Dear Editor and Reviewers,

We sincerely thank you for your careful evaluation of our manuscript and for the constructive comments and valuable suggestions. We have carefully considered all comments and revised the manuscript accordingly. We believe that these revisions have substantially improved the quality, clarity, and scientific rigor of the manuscript.

Our point-by-point responses to each comment are provided below. Reviewer comments are presented in bold, our responses are shown in regular text, and all corresponding revisions in the revised manuscript are highlighted for ease of reference.

We appreciate your time and consideration and hope that the revised manuscript meets the requirements for publication.

(1) The revised methodology is not reflected in the public dataset, which violates basic expectations for data intensive journals. Substantial changes to the data generation method would logically require corresponding updates to the final dataset, yet neither the manuscript nor the public data repository has been updated to reflect these methodological revisions. The authors introduced important methodological improvements, including dual threshold segment fusion for natural and planted forests, correction of systematic age underestimation, labeling of old growth forests as ≥40 years, a pixel level Spatial Uncertainty Index (SUI), expanded validation samples, and quantitative comparison with the GAMI dataset. However, the archived dataset still corresponds to the original version published on November 25, 2025. It does not include the SUI layer, segment fusion, dual threshold adjustment, or ≥40 year age classification, and the metadata still describe the original CCDC breakpoint method. This discrepancy between the manuscript and the publicly available data is highly problematic for a data journal, as it prevents reproducibility and could mislead future users.

Thank you for this important comment. We fully agree that the archived dataset should be fully consistent with the methodology described in the revised manuscript.

In response, we have regenerated the dataset using the final processing workflow and deposited this revised version in Figshare under a new permanent DOI: 10.6084/m9.figshare.33005915

As the manuscript is still under peer review, Figshare keeps the deposited dataset private until formal acceptance. To enable verification during the review process, we provide the following private review link for editors and reviewers: https://figshare.com/s/93ad7c364df2e67c7339

Upon acceptance, the dataset will be released publicly through the registered DOI, ensuring full consistency between the published manuscript and the archived data. We have also revised the Data Availability section to clarify the dataset status and its accessibility during peer review.

(2) CCDC cannot distinguish non stand replacing disturbances, leading to systematic underestimation of forest age. The CCDC algorithm detects both stand replacing and non stand replacing disturbances but cannot separate them. Partial disturbances including selective logging, thinning, insect outbreaks, low severity fire, wind damage, or understory removal can trigger a breakpoint and reset the pixel's forest age to zero even when most

overstory trees remain. For instance, selective removal of 20% of a 100 year old stand would lead CCDC to assign a near zero age, even though 80% of the trees remain over 100 years old. This limitation causes widespread, systematic underestimation of forest age and reduces the reliability of the dataset for carbon accounting, forest dynamics, and ecological modeling. The authors have not adequately addressed this issue.

We sincerely thank the reviewer for this insightful comment regarding the potential impact of non-stand-replacing disturbances on CCDC-derived forest age estimates. We agree that partial disturbances, including selective logging, thinning, insect outbreaks, low-severity fires, and wind damage, may generate detectable spectral breakpoints and consequently lead to age underestimation when interpreted as stand-replacing disturbances.

In response to this concern, we have revised the manuscript to clarify the conceptual definition and limitations of CCDC-derived forest age. Specifically, we now explicitly define forest age in this study as disturbance-recovery age derived from detectable spectral changes in Landsat time-series observations, rather than inventory-based stand age or the biological age of surviving trees. We also clarify that CCDC detects spectral disturbances but cannot distinguish between stand-replacing and non-stand-replacing disturbance events.

We have further clarified that the segment-fusion strategy was not designed to identify or correct non-stand-replacing disturbances. Instead, this procedure aims to reduce artificial fragmentation of temporally continuous recovery trajectories caused by spectral segmentation during the transition from rapid post-disturbance spectral recovery to spectral stabilization.

We acknowledge that partial disturbances remain an inherent uncertainty source in optical time-series-based forest age estimation. Such events may cause localized underestimation because the detected breakpoint represents the timing of spectral change rather than complete stand replacement. The magnitude of this uncertainty varies spatially depending on disturbance frequency, intensity, and forest management practices. Therefore, GFAM-N/P should be interpreted as a disturbance-recovery age product and applied cautiously in studies requiring true stand age information, such as forest demographic analyses or carbon stock initialization.

These clarifications have been incorporated into Section 2.4.1 (Definition of forest age), Section 2.4.2 (Segment-fusion logic), and Section 4.3 (Systematic biases and limitations).

(3) Fundamental concerns with the CCDC algorithm and parameterization remain. The authors added sensitivity analysis for key parameters but still apply a globally uniform set of thresholds even though tropical, temperate, and boreal regions likely require different configurations. The slope thresholds of 0.03 and 0.04 used for segment fusion are not transparently derived or justified. The t_start lag effect in the GEE CCDC/COLD implementation still introduces systematic underestimation of forest age, and the proposed segment fusion procedure is presented as an ad hoc post processing step without fully reproducible documentation.

We sincerely thank the reviewer for this insightful comment regarding the robustness, transparency, and reproducibility of the CCDC-based age estimation framework. In the revised manuscript, we substantially strengthened the methodological description, parameter justification, sensitivity evaluation, and discussion of algorithmic limitations to address these concerns.

First, to evaluate the suitability of a globally consistent parameterization scheme, we conducted a comprehensive sensitivity analysis of the two key CCDC parameters (minObservations and chiSquareProbability) across four major Köppen climate zones (Tropical, Temperate, Boreal, and Arid) (Section 3.2; Fig. 8). The results indicate that although some regional differences exist in parameter sensitivity, the selected configuration (minObservations = 6; chiSquareProbability = 0.99) consistently achieves robust performance across all climate zones. We further discuss in Section 4.1 that although regionally adaptive parameterization may better capture local variations in forest recovery dynamics, it may also introduce artificial spatial discontinuities and reduce cross-regional comparability. Therefore, a globally standardized parameter framework was intentionally adopted to ensure methodological consistency and comparability of the final global product.

Second, the derivation of the segment-fusion thresholds has now been explicitly documented. In the newly added Section 3.1.1 (“ Spectral successional trajectories and calibration of fusion thresholds ”), we analyzed post-disturbance recovery trajectories using a globally distributed reference database covering tropical, temperate, boreal, and arid ecosystems. The results show that recovery slopes of natural forests generally cluster between 0.025 and 0.033 (global mean = 0.029), whereas planted forests cluster between 0.035 and 0.044 (global mean = 0.038). Based on these empirically derived distributions, generalized thresholds of 0.03 for natural forests and 0.04 for planted forests were adopted. Figure 3 has been added to document this calibration process and the empirical basis for the selected thresholds.

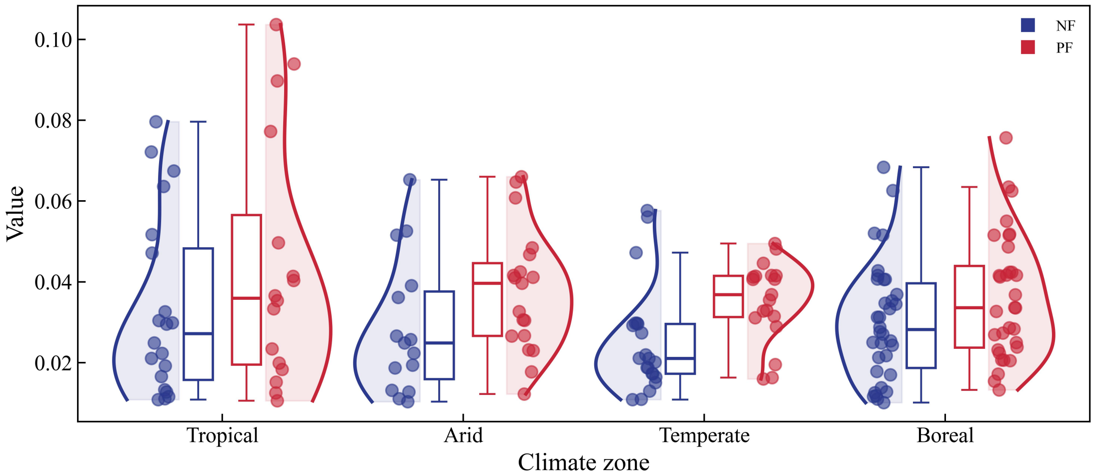

Figure 3. Statistical distributions of the post-disturbance spectral recovery slopes for NF and PF across major global climate zones.

Third, we acknowledge that the tStart variable in the GEE implementation of CCDC reflects model stabilization rather than the exact biological onset of forest recovery. To explicitly address this issue, we added a dedicated discussion in Section 4.3. As explained in the revised manuscript, this mechanism introduces a temporal lag whose magnitude depends on regional observation frequency and the availability of valid Landsat observations. The segment-fusion framework was designed to reduce the age underestimation associated with delayed identification of recovery trajectories by linking temporally continuous recovery segments and tracing them back toward the onset of post-disturbance spectral recovery. However, this procedure cannot completely eliminate the inherent initialization lag of the CCDC model, and the remaining uncertainty associated with model initialization is now explicitly discussed in Section 4.3.

Finally, we clarify that the segment-fusion procedure is not an ad hoc post-processing operation. To improve reproducibility, Section 2.4.2 has been expanded to provide a complete description of the fusion workflow. The procedure is fully deterministic and governed by four explicit criteria: (1) identification of a terminal stable segment (|slope| < 0.01), (2) detection of a preceding recovery segment, (3) forest-type classification (natural versus planted forest), and (4) application of the corresponding empirically calibrated recovery threshold (0.03 or 0.04). No manual intervention or region-specific adjustment is applied during processing.

Accordingly, the revised manuscript now provides a more transparent and reproducible description of the CCDC parameterization and segment-fusion framework. The workflow can be reproduced directly from the methodological description without manual intervention or region-specific tuning, while the remaining limitations of time-series-based forest age estimation are explicitly acknowledged.

(4) Validation samples are spatially imbalanced and lack global representativeness. The global validation sample map (Figure 3a) shows strong spatial bias. Regions including China, East Asia, and Southeast Asia are dominated by natural forest (NF) samples with very few planted forest (PF) samples. Sample density and representation are inconsistent across continents. Such an unbalanced and non stratified sampling design violates standard principles for validating large scale remote sensing products. As a result, the reported overall accuracy, RMSE, and R^2 may not be globally representative, and the validation does not fully support the claim of a reliable global forest age dataset.

We thank the reviewer for raising this important concern regarding the representativeness of the validation dataset. Contrary to a simple random sampling strategy, the validation samples in this study were generated using a globally stratified sampling framework designed to achieve broad spatial coverage and balanced representation of forest types.

As described in Section 2.5, the global land surface was first partitioned into 57,559 independent 0.5° × 0.5° grid cells. Within each grid, a stratified random sampling strategy was applied to select samples separately for natural forests (NF) and planted forests (PF), thereby ensuring broad spatial coverage while minimizing sample clustering. The final reference dataset contains 8,085 samples, including 4,021 NF samples and 4,064 PF samples, indicating a nearly balanced representation of the two forest types rather than a dominance of NF samples.

The apparent imbalance of sample distributions in Figure 4a may partly result from visualization limitations caused by point overlap at the global display scale. Because thousands of validation samples are displayed simultaneously, some PF samples located within mixed forest landscapes may be visually obscured by neighboring NF samples. To improve interpretation, Figure 4a has been redrawn with enhanced visualization of sample locations.

Furthermore, variations in sample density among continents primarily reflect differences in the spatial extent and distribution of global forests rather than deficiencies in the sampling framework. Regions with extensive forest coverage naturally contain more validation samples, whereas areas with limited forest distribution contribute fewer samples. This design follows the principle of proportional representation while maintaining global spatial coverage. Additional clarification regarding the sampling strategy and sample composition has been added to Section 2.5.

Therefore, based on the globally stratified sampling framework, balanced NF/PF representation, and broad geographic coverage, the reference dataset provides an appropriate basis for evaluating the performance of the global disturbance-recovery age product.

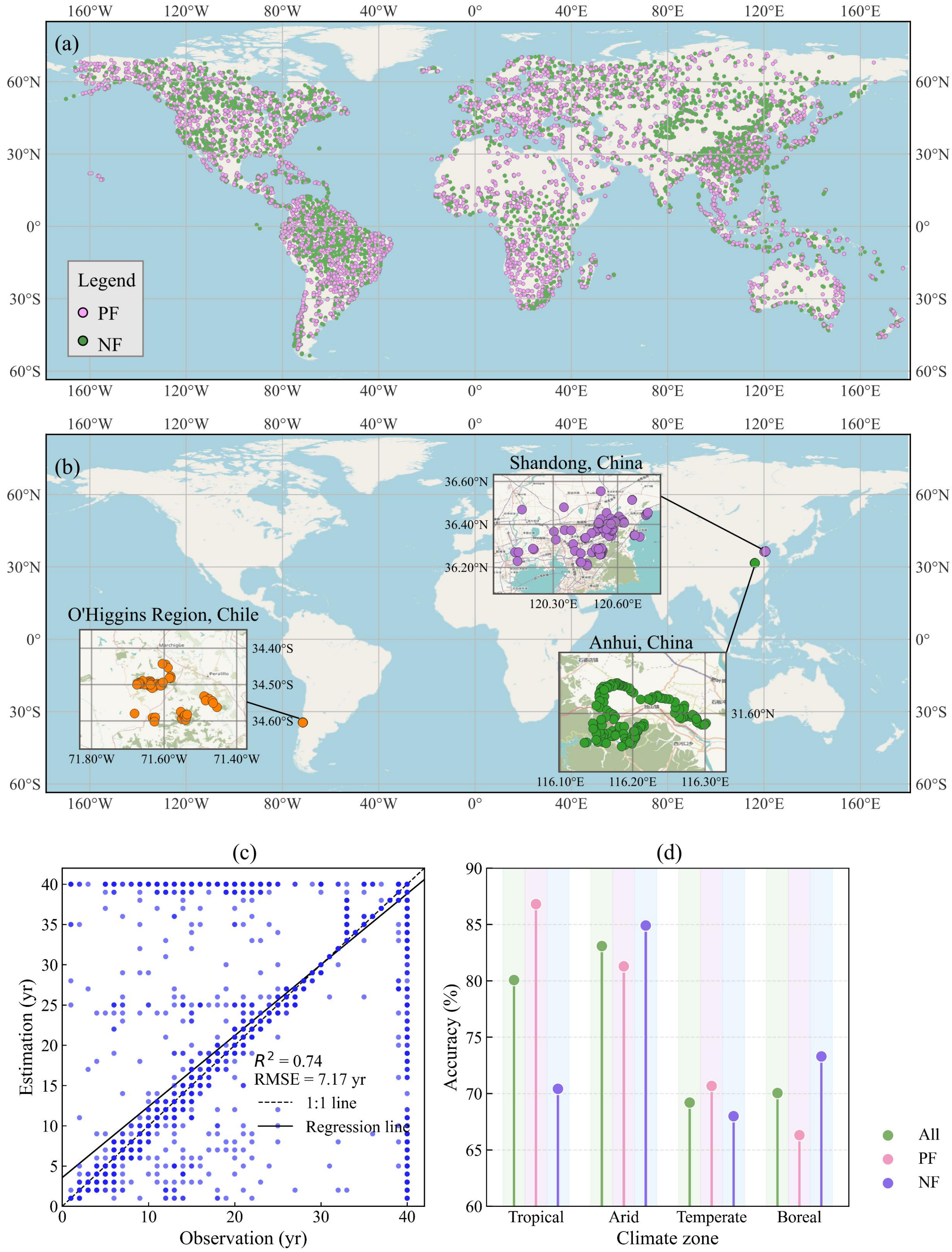

Figure 4: Spatial distribution and validation of reference samples for global forest age assessment. (a) The global distribution of all validation samples. (b) The global distribution of field samples. (c) Scatter plot and fitting relationship between observed and predicted forest ages. (d) Accuracy comparison across climate zones and sample types. NF represents natural forest, PF represents planted forest, and All represents the combination of the two types of forests.

(5) Threshold sensitivity relies on OA and overlooks imbalanced PA and UA. The sensitivity analysis (Figure 8) uses OA as the only metric, despite known limitations in disturbance and forest age mapping. OA tends to be dominated by large stable non forest areas and can mask imbalances in PA and UA. The analysis does not reveal how parameters influence omission and commission errors or how they affect systematic overestimation or underestimation of forest age. A more complete evaluation would be needed to justify the chosen parameter set.

We thank the reviewer for this insightful comment regarding the limitations of relying solely on overall accuracy (OA) in the sensitivity analysis. In response, we have revised Section 3.2 (Sensitivity of CCDC configurations to spectral inputs and algorithmic parameters) by incorporating producer's accuracy (PA) and user's accuracy (UA), which provide a more comprehensive assessment of omission and commission errors associated with different parameter configurations.

Specifically, in the revised Section 3.2, the evaluation framework has been expanded beyond OA to include class-wise PA and UA derived from age-stratified confusion matrices. The forest age was consistently grouped into five types (1–10, 10–20, 20–30, 30–39, and ≥40 years) across all climate zones to ensure comparability of accuracy metrics. Figure 11 has been added.

The results show that increasing minObservations from 5 to 6 leads to a consistent improvement in both PA and UA across all climate zones, followed by a clear saturation effect beyond this threshold. Similarly, the chiSquareProbability parameter exhibits a clear transition from unstable low-accuracy regimes (0.7–0.8), characterized by increased omission and commission errors, to a stable high-accuracy regime at 0.99, where both PA and UA are substantially improved and inter-class variability is significantly reduced. We have updated the revisions in Section 3.2.

In addition, because the sensitivity analysis was conducted using categorical forest age classes rather than continuous age estimates, PA and UA were used to evaluate class-specific omission and commission errors under different parameter configurations. potential systematic biases in continuous age estimation, including underestimation caused by partial disturbances and delayed breakpoint initialization, as well as discrepancies between disturbance-recovery age and actual stand establishment time due to ecological recovery lags, are further discussed in Section 4.3.

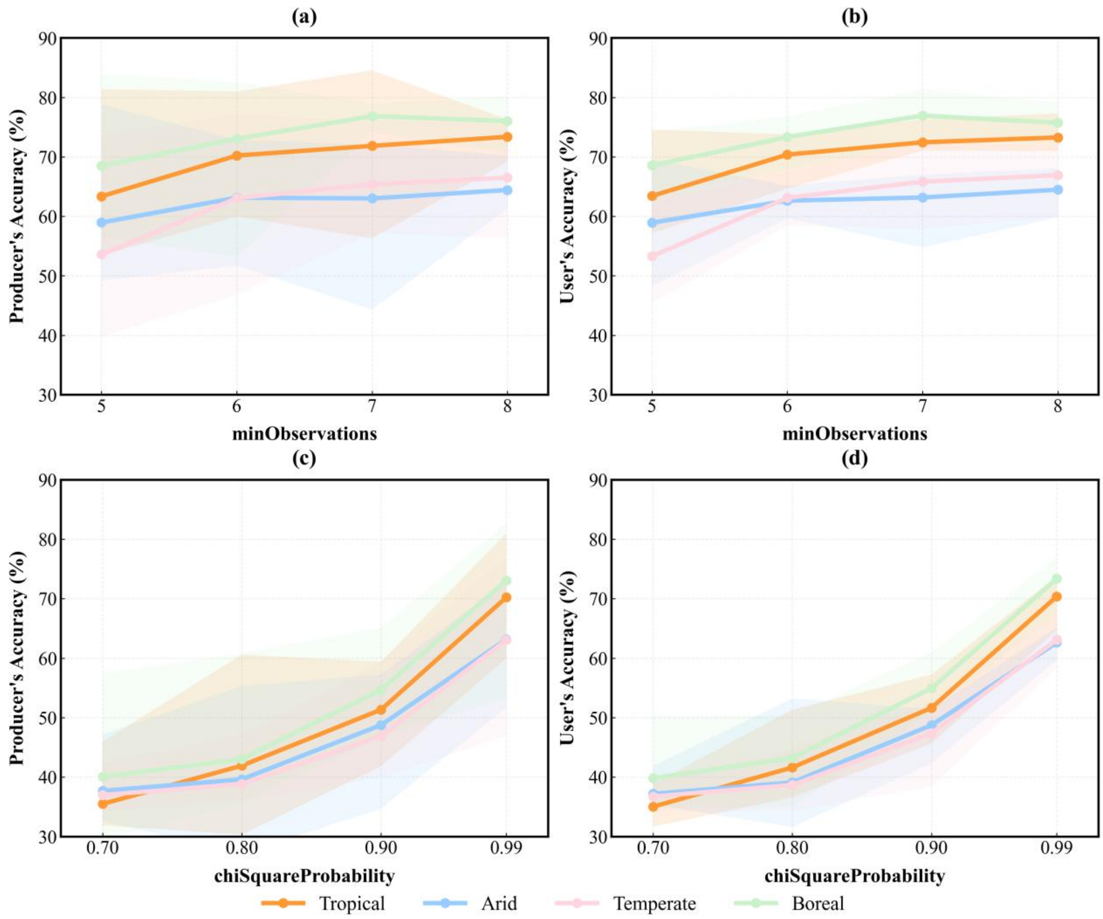

Figure 11: Response of PA and UA to variations in (a, b) minObservations and (c, d) chiSquareProbability across four major climate zones. Shaded areas denote the [Min, Max] uncertainty intervals.

(6) The validation framework remains unreliable and insufficiently global. The primary validation continues to depend on visual interpretation from Google Earth, which is limited by cloud cover, missing historical imagery, and inconsistent interpretation. Although the authors added validation using the US FIA dataset in Section 3.1.4, this approach has important limitations. FIA represents only temperate forests of the United States and does not reflect global biomes. The definition of stand age in FIA differs conceptually from disturbance–recovery age in this study, and spatial offsets in FIA plot coordinates prevent precise matching with 30 m pixels. Validating a global dataset using only one national inventory is methodologically challenging and does not meet the expectations of top remote sensing journals.

We sincerely thank the reviewer for this valuable comment regarding the representativeness and reliability of the validation framework. We agree that validating a global disturbance-recovery age product using only a single national inventory dataset (e.g., US FIA) would be insufficient to demonstrate global applicability, particularly because national inventories differ in forest definitions, age concepts, and spatial sampling strategies, which may affect direct comparisons with satellite-derived disturbance–recovery age estimates. In the revised manuscript, we have clarified the limitations of individual validation sources and substantially strengthened the validation framework by integrating multiple independent sources of evidence.

Specifically, we expanded and clarified the validation framework as follows:

(1) Global stratified interpretation samples. We retained the globally distributed reference dataset consisting of 8,085 samples (4,021 NF samples and 4,064 PF samples), which were generated using a globally stratified sampling framework based on 57,559 independent 0.5° × 0.5° grid cells. Importantly, these reference samples were not derived from Google Earth imagery alone. Instead, disturbance-recovery age labels were established using a multi-evidence interpretation protocol integrating annual Landsat spectral trajectories (NBR and NDVI), Global Forest Change (GFC) disturbance records, and historical high-resolution imagery from Google Earth. In this framework, GFC loss years and Landsat-derived spectral breakpoints were jointly used to identify the timing of the most recent detectable stand-replacing disturbance, while historical imagery was used as supplementary spatial evidence for verifying canopy conditions and disturbance patterns. Samples with ambiguous disturbance histories, insufficient historical observations, or inconsistent evidence among data sources were excluded through strict confidence screening. To reduce interpreter-related inconsistencies, only samples supported by consistent evidence from multiple sources were retained. This procedure has been further clarified in Section 2.5.

(2) Independent forest inventory validation. We retained independent forest inventory-based validations from the United States, China, and Chile. We acknowledge that national inventory datasets have inherent limitations, including regional coverage constraints and differences between inventory-based stand age and disturbance–recovery age estimated from satellite time series. Therefore, these datasets were not used as the sole basis for global validation, but rather as independent regional benchmarks to evaluate the consistency between satellite-derived age estimates and inventory observations under different forest conditions.

(3) Independent regional product comparison. To further enhance the global evaluation, we added an independent comparison with the recently published 30-m Canadian forest age product (Maltman et al., 2023), which is described separately in Section 3.1.3. Canada was partitioned into 100-km grids, and stratified random sampling was conducted across five forest age classes. After removing invalid pixels, 11,092 paired samples were retained. The comparison achieved an overall agreement of 72.23%, with particularly strong agreement for mature forests (PA = 83.44%; UA = 81.50%). Although the Canadian product adopts a different methodology by integrating disturbance mapping, spectral recovery, and allometric information, the agreement between two independently developed approaches provides additional evidence for the robustness of the GFAM-N/P dataset.

We have also revised the manuscript to explicitly acknowledge that no globally harmonized forest inventory currently exists for forest age validation. Therefore, we adopted a multi-source validation strategy combining globally distributed reference samples, national forest inventories, and independent regional forest age products. This integrated framework provides complementary evidence across different spatial scales and methodological perspectives, while avoiding reliance on any single validation source.

(7) The separation between natural and planted forests remains a post hoc procedure. The authors present their product as the first global forest age dataset to distinguish natural and planted forests, but the separation relies on an external GNPF mask applied after processing. The core age estimation algorithm is identical for both forest types, without truly differentiated modeling tailored to their different disturbance and growth characteristics.

The dual threshold fusion is a post processing step rather than a genuine methodological innovation. Therefore, the claimed advance remains partially overstated.

We agree that the separation between natural forests (NF) and planted forests (PF) in our framework is not achieved through independent age-estimation models, but rather through a forest-type-informed stratification procedure applied during the post-segmentation refinement stage. Therefore, we acknowledge that our original wording may have overstated the degree of methodological differentiation between forest types.

Our intention was not to claim that the underlying CCDC age-estimation algorithm is fundamentally different for NF and PF. The core temporal segmentation and age estimation procedures remain identical for both forest types. The distinction is introduced after segmentation through the integration of the Global Natural and Planted Forest (GNPF) dataset, which provides forest-origin information used to guide forest-type-specific segment refinement and age adjustment. Specifically, different segment-fusion thresholds (>0.03 for NF and >0.04 for PF) were applied to account for their contrasting recovery characteristics during post-disturbance succession.

To avoid overstating the novelty of this component, we have revised the manuscript throughout the Abstract, Introduction, and Conclusion. We no longer describe the approach as a fundamentally different age-estimation model for NF and PF. Instead, we consistently refer to it as a forest-type-informed stratification strategy, in which forest-origin information is incorporated during the post-segmentation refinement stage. We have also revised the discussion of previous products to emphasize that most existing forest age datasets estimate stand age without explicitly representing forest origin, rather than implying that they employ comparable stratification procedures.

Our primary contribution is therefore clarified as the development of a globally consistent 30-m forest age dataset that provides disturbance-recovery age information separately for natural and planted forests within a unified age-estimation framework, rather than the introduction of two fully independent forest-type-specific age models.

Sincerely,

The Authors