Dear Topical Editor and Reviewers:

On behalf of my co-authors, we sincerely thank you for the second round of reviews and for the additional insightful comments on our manuscript. We greatly appreciate your continued guidance in refining our study, entitled “1 km annual forest cover and plant functional type dataset for China from 1981 to 2023” (essd-2025-475).

We have further revised the manuscript thoroughly to address the remaining concerns. A detailed, point-by-point response to the new set of comments is provided below.

Looking forward to hearing from you soon.

Best regards,

Bo Liu

School of Geography and Tourism, Shaanxi Normal University

No. 620, West Chang’an Street, Chang’an District, Xi’an 710119, China

Response to comments

Paper: essd-2025-475

Title: 1 km annual forest cover and plant functional type dataset for China from 1981 to 2023

Journal: Earth System Science Data

Response to the Reviewers #1

RC1: Comment on essd-2025-475

[Response to major comment 2] You replaced the NDVI dataset with the Landsat-based composite from Cai et al. (2025) and used the 1985 data as a proxy for 1981-1984 due to a lack of imagery. Could you provide a quantitative assessment or a sensitivity analysis demonstrating that the ecological conditions (e.g., climate, vegetation phenology) in 1985 are a sufficiently representative proxy for the 1981-1984 period?

Response: We thank the reviewer for this constructive suggestion. To quantitatively evaluate the validity of this substitution strategy, we used the 1985 dataset from Cai et al. (2025) as a baseline and conducted a comprehensive pixel-wise correlation analysis against three independent long-term NDVI products covering the 1981–1985 period: SNU NDVI (1982–1985; Jeong et al., 2024), PKU GIMMS NDVI (1982–1985; Li et al., 2023), and Li NDVI (1981–1984; Li et al., 2024).

To ensure comparability, all validation datasets were processed to derive the growing season (June–October) median NDVI, consistent with the compositing method used by Cai et al. (2025). Furthermore, prior to analysis, the Cai et al. dataset was resampled using bilinear interpolation to match the spatial resolution of each respective NDVI product (SNU: 0.05°; PKU GIMMS: 0.083°; Li: 0.05°).

The results demonstrate that the 1985 data from Cai et al. (2025) exhibits consistent and significant positive correlations with the annual data of all validation products throughout the 1981–1985 period (see **Figure S1**). The coefficient of determination (R^2) consistently exceeded 0.8, indicating that the spatial pattern of vegetation density in 1985 is highly representative of the entire 1981–1984 period. Crucially, the correlation metrics exhibited remarkable temporal stability. For any given validation product, the R^2 values derived from comparing its annual data (1981–1985) against the Cai et al. 1985 proxy remained highly consistent, with a maximum intra-group R^2 difference of less than 0.04. This suggests that interannual fluctuations in vegetation state driven by climate or phenology were minimal within this specific 5-year window.

Cross validation using these multisource datasets further confirms that the Landsat based 1985 composite successfully captures the fundamental ecological characteristics of the early 1980s. Therefore, despite minor interannual variations, we consider the 1985 dataset a robust and reliable proxy for the 1981–1984 period in the absence of earlier Landsat imagery.

We have added the above explanation to **Section 2.3, Lines 151–154** of the revised manuscript to quantitatively justify the gap-filling strategy.

"To validate this substitution, we performed a pixel wise correlation analysis between the 1985 growing season NDVI and three independent, coarser resolution products covering 1981–1985 (Jeong et al., 2024; Li et al., 2023; Li et al, 2024). The results indicate that the 1985 dataset maintained high spatial consistency with the validation products ($R^2 > 0.8$). Furthermore, the correlation metrics exhibited exceptional temporal stability ($\Delta R^2 < 0.04$; Fig. S1)."

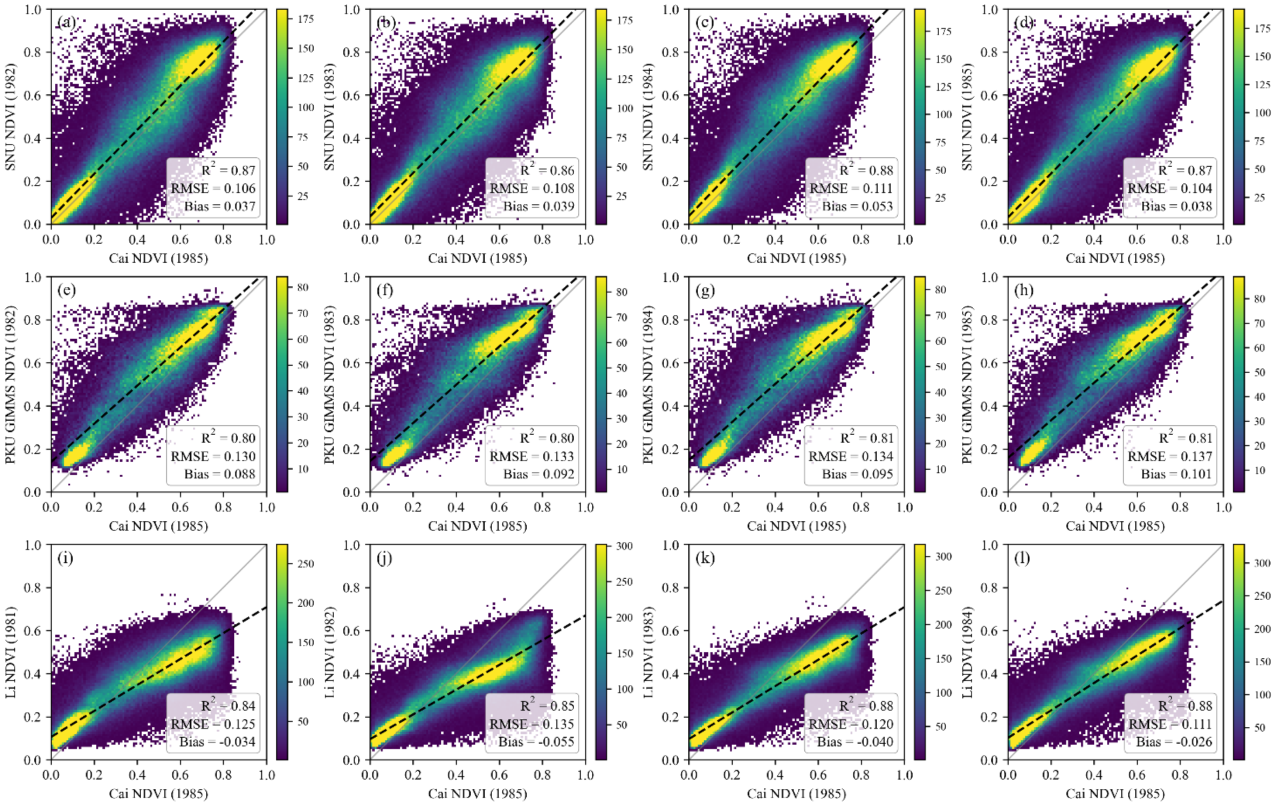

Figure S1. *Pixel wise correlation analysis validating the representativeness of the 1985 NDVI dataset for the early 1980s. The scatter plots compare the growing season median NDVI from Cai et al. (2025) (1985 baseline, x-axis) against concurrent and preceding annual data from three independent long-term products (y-axis): (a–d) SNU NDVI (1982–1985), (e–h) PKU GIMMS NDVI (1982–*

1985), and (i–l) Li NDVI (1981–1984). The color scale represents point density. The black dashed line indicates the 1:1 line, and the solid grey line represents the linear regression fit. Statistical metrics (R^2, RMSE, and Bias) are provided in each panel.

References:

Cai, Y., Li, X., Zhu, P., Nie, S., Wang, C., Liu, X., and Chen, Y.: China Earth Observation Data Cube: The 30-m Seamless Annual Leaf-On Landsat Composites from 1985 to 2023, *J. Remote Sens.*, 5, 0698, https://doi.org/10.34133/remotesensing.0698, 2025.

Jeong, S., Ryu, Y., Gentine, P., Lian, X., Fang, J., Li, X., Dechant, B., Kong, J., Choi, W., and Jiang, C.: Persistent global greening over the last four decades using novel long-term vegetation index data with enhanced temporal consistency, *Remote Sens. Environ.*, 311, 114282, https://doi.org/10.1016/j.rse.2024.114282, 2024.

Li, H., Cao, Y., Xiao, J., Zhao, Y., and Li, X.: A daily gap-free normalized difference vegetation index dataset from 1981 to 2023 in China, *Sci. Data*, 11, 527, https://doi.org/10.1038/s41597-024-03364-3, 2024.

Li, M., Cao, S., Zhu, Z., Wang, Z., Myneni, R. B., and Piao, S.: Spatiotemporally consistent global dataset of the GIMMS Normalized Difference Vegetation Index (PKU GIMMS NDVI) from 1982 to 2022, *Earth Syst. Sci. Data*, 15, 4181–4203, https://doi.org/10.5194/essd-15-4181-2023, 2023.

[Response to major comment 5] Your revised validation uses LandTrendr on Landsat series to identify "stable" NFI plots and then backcasts labels. For pixels identified as "unstable" (changed), how did you assign a historical PFT label prior to the change event? Could you elaborate on the specific criteria and the number of experts involved in the manual interpretation to ensure the objectivity and reproducibility of this historical label assignment process?

Response: We appreciate your feedback. To address the processing of "unstable" pixels and the assignment of historical labels, we employed a cross-validation strategy to pinpoint the timing of land cover transitions. Specifically, we integrated raw NBR values derived from Landsat time series, LandTrendr-fitted temporal trajectories, and high-resolution historical imagery time series from Google Earth (Figs. S8 and S9). Our analysis focused on the segments exhibiting the greatest magnitude of change between LandTrendr-fitted vertices, examining the high-resolution imagery features immediately preceding and following these events.

Through visual interpretation of spectral variations (e.g., changes in vegetation greenness) and textural features (e.g., exposed soil following deforestation or canopy texture post-afforestation) in Google Earth imagery, we identified the exact year of land cover conversion. By combining this with LandTrendr-fitted NBR trajectories, we assigned a forest/nonforest status to each individual year (e.g., in Fig. S8, the plot is classified as forest from 1985–2012 and nonforest from 2013–2023). Once the specific year of the change event was determined, the PFT label obtained from the NFI field survey was consistently assigned to all years within the 1985–2023 period where the plot was identified as "forest." This labelling approach is based on our underlying hypothesis that PFT distributions are primarily controlled by relatively stable long-term climatic, edaphic, and topographic conditions. Consequently, we assumed that these distributions remained consistent throughout the 1981–2023 period.

For the early period of 1981–1984, due to the scarcity of Landsat and high-resolution imagery, we employed an extrapolation approach based on the plot's initial state in 1985. Specifically, if a plot was identified as a specific forest PFT or nonforest in 1985, we assumed it retained the same attribute throughout 1981–1984. This approach assumes that reversals between forest and nonforest states (e.g., forest–nonforest–forest transitions) are unlikely within a five-year window.

To ensure the objectivity of this visual interpretation process, ten experts with extensive experience in remote sensing data analysis were invited to participate in the validation. For samples with ambiguous features or those difficult to classify, a team discussion was initiated, and the final reference labels were assigned only after reaching a majority consensus.

Ultimately, by integrating the annual records from both stable plots and the unstable plots verified through this procedure, we constructed a comprehensive validation database covering the entire 1981–2023 period.

We have incorporated these relevant details into the revised manuscript, specifically in **Section 3.5, Lines 399–412**.

"For unstable plots, we employed a cross-validation strategy to pinpoint the precise timing of land cover transitions. Specifically, we integrated raw NBR values derived from Landsat time series, LandTrendr-fitted temporal trajectories, and historical high-resolution imagery from Google Earth. We paid particular attention to validating the image features corresponding to the segments with the

greatest magnitude of change between fitted vertices (Figs. S8 and S9). Through visual interpretation of spectral variations and textural features in the Google Earth imagery, we identified the exact year of land cover conversion (e.g., forestland converted to nonforest in 2013; Fig. S8). Consequently, the specific PFT label for each plot was consistently applied to all years within the 1985–2023 period where the plot was identified as forestland.

For the 1981–1984 period, we extrapolated the classification status from 1985 to the preceding years (i.e., plots identified as a specific forest PFT or nonforest in 1985 were assigned the same attribute for 1981–1984). This approach assumes that reversals between forest and nonforest states (e.g., forest–nonforest–forest transitions) are unlikely to occur within a five-year window. The labelling was conducted by ten experts experienced in remote sensing data analysis. To ensure rigorous quality control, the samples with ambiguous features were subjected to team discussion, and the final reference labels were assigned only after a majority consensus. Finally, by integrating the annual records from both stable and unstable plots, we constructed a validation database spanning the entire 1981–2023 period."

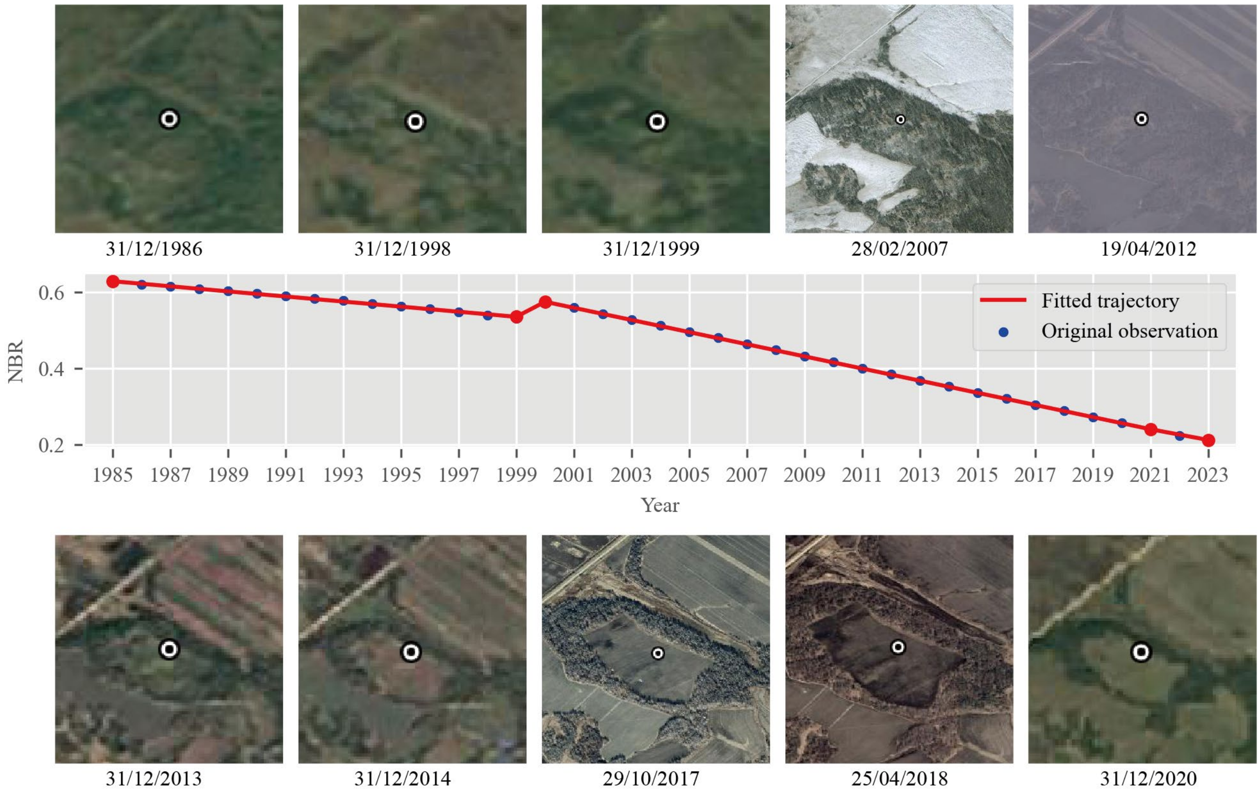

Figure S8. *Time series example of an unstable plot (forest to nonforest transition). A forest loss event occurred in 2013. The historical Google Earth imagery time series illustrates the conversion process from forestland to nonforest. (Example location: 48.066463°N,* 127.562063 °E). In the time series plot, the blue points represent Landsat observations, the red lines denote the LandTrendr model fits, and the red dots indicate the vertices identified by the LandTrendr model. The concentric circles mark the location of the sample plot.*

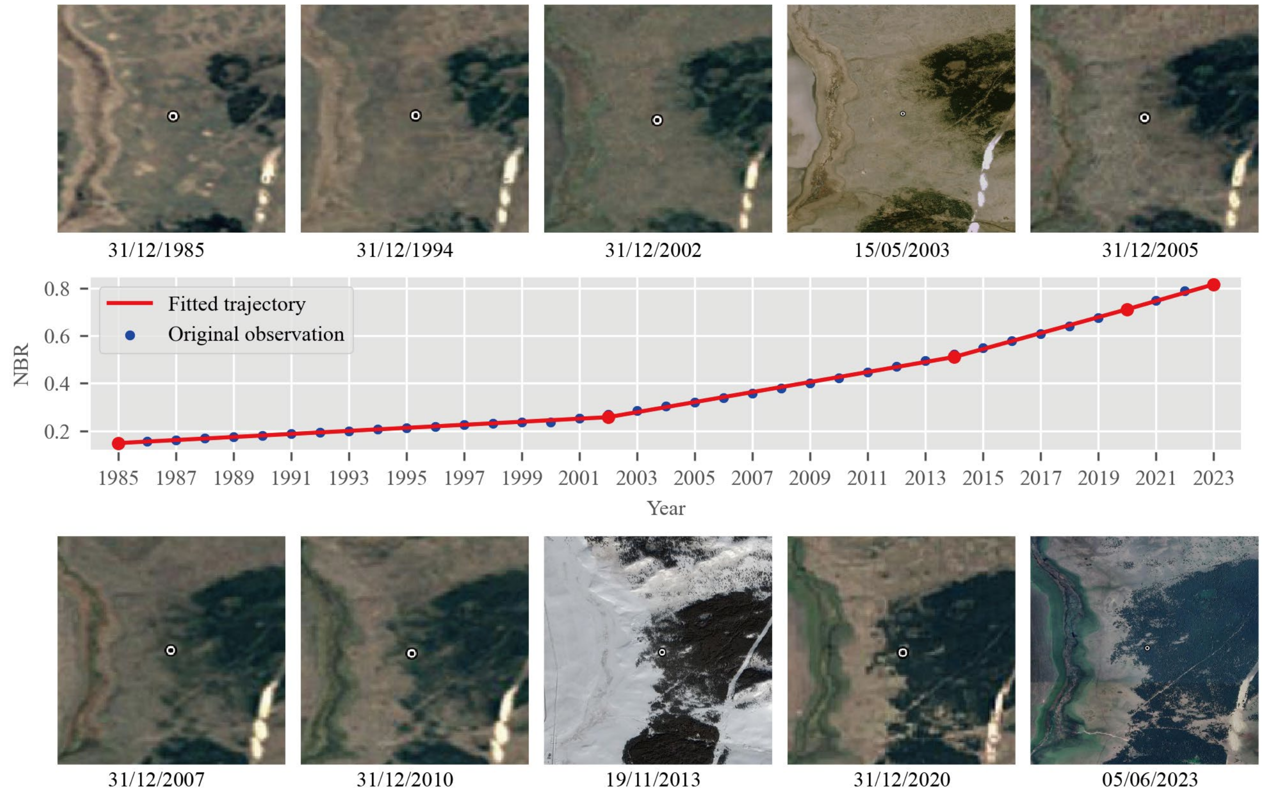

Figure S9. *Time series example of an unstable plot (nonforest to forest transition). A forest gain event occurred in 2007. The historical Google Earth imagery time series illustrates the conversion process from nonforest to forest. (Example location: 47.857612°N, 119.494846°E). In the time series plot, the blue points represent Landsat observations, the red lines denote the LandTrendr model fits, and the red dots indicate the vertices identified by the LandTrendr model. The concentric circles mark the location of the sample plot.*

[Response to major comment 6] Adopting the "wooded land" definition (excluding shrubland) is a key methodological improvement. However, could this stricter definition potentially lead to an underestimation of forest cover in arid/semi-arid regions where forests may naturally have lower canopy cover or are sparse?

Response: We appreciate your insightful comments and respond to your concerns about the following three aspects:

First, regarding the estimation of total forest area in arid/semi-arid regions: We confirm that the strict exclusion of shrubland does not lead to an underestimation of total forest area (within the context of our study's definition) in arid or semi-arid regions. This is attributed to the "forest consistency" algorithm employed in this study, which dynamically adjusts the consistency threshold at the provincial level until the reconstructed forest area precisely matches the corresponding NFI statistical values. Consequently, even if certain satellite products exhibit lower sensitivity to sparse forests in these regions (resulting in lower consistency scores for such pixels), our method automatically lowers the threshold to force the inclusion of these low-consistency pixels to fill the area gap. This mechanism ensures that the reconstructed total area remains strictly faithful to NFI ground-based inventory data, thereby eliminating the risk of underestimation caused by weak spectral signals from sparse forests.

Second, regarding spatial distribution and scale effects: We acknowledge that our product may omit the spatial distribution of some sparse forest pixels. However, this is not a consequence of our stricter definition, but rather an inherent "scale aggregation effect" of binary rasters at a 1km resolution. At the 1km scale, denser forest centers are more likely to be retained, while sparse and dispersed marginal forest pixels are more prone to omission.

While incorporating shrubland might appear to be a potential strategy to recover these omitted pixels, it risks conflating categories that are strictly distinct in the NFI definitions. NFI data indicates that in Northwest China, the area of shrubland is comparable to, or even exceeds, that of wooded land (For instance, data from the Seventh National Forest Inventory (NFI 7) indicates that in Xinjiang, Qinghai, and Ningxia, the total area of wooded land amounted to approximately 2.55 million hectares (Mha), whereas the shrubland area was substantially larger, reaching 7.87 Mha. (Detailed statistics are available at the National Forestry and Grassland Science Data Center (https://www.forestdata.cn/)). Therefore, including shrubland area would introduce commission errors, causing large areas of shrubland to be misclassified as forestland. Faced with the trade-off between "capturing core forest areas while missing marginal pixels" and "including sparse forest but introducing misclassification," we prioritized the former. While this limitation of binary classification could be addressed in the future through soft classification using sub-pixel fractional cover (i.e., reporting the proportion of land cover types within a pixel rather than a categorical class), the 1km resolution provided by this dataset is sufficient to characterize the spatial heterogeneity required by current Dynamic Global Vegetation Models (DGVMs), which typically operate at grid resolutions of 0.1° to 0.5°.

Finally, regarding the biophysical justification for DGVMs: We emphasize that the strict distinction between trees and shrubs is driven by the necessity to provide accurate Plant Functional Type (PFT) maps for DGVMs. In most dynamic vegetation models, "trees" and "shrubs" represent distinct PFTs with fundamentally different carbon allocation strategies (Verbruggen et al., 2021; Yu et al., 2022). Shrubs are typically constrained in height and canopy expansion by specific allometric scaling relationships, and their post-harvest carbon pool turnover patterns more closely resemble those of herbaceous PFTs. Merging shrubland into the forest category would cause models to overlook these differences in canopy structural constraints and carbon construction costs. This, in turn, would lead to the misrepresentation of the light competition environment and vertical Leaf Area Index (LAI) distribution within vegetation communities, introducing significant uncertainty into the simulation of key surface fluxes (e.g., GPP and ET). Therefore, our approach ensures a rigorous biophysical distinction between tree PFTs and shrub/grass PFTs, which is critical for improving the simulation accuracy of dynamic vegetation models.

We have revised the relevant description in the Discussion section to improve precision. Specifically, we have updated the sentence *"By excluding shrublands, our extracted distribution of 'potential forest' aligns more closely with the actual forest distribution in sparsely forested regions such as Xinjiang and Qinghai"* to read: *"By excluding shrublands, our extracted distribution of 'potential forest' aligns more closely with the* ***core*** *forest distribution in regions such as Xinjiang and Qinghai."*

Furthermore, we have added a new paragraph to explicitly discuss the limitations associated with the 1 km resolution of the product and to outline potential directions for future improvements. This addition can be found in the discussion section (Lines 706–715):

"Second, our method employs a 1 km binary classification scheme. This approach may induce scale aggregation effects in arid and semiarid regions where forest cover is naturally sparse and fragmented. Specifically, dispersed forest pixels risk being omitted if they do not constitute the dominant land cover within a 1 km grid cell. While incorporating shrublands might capture the spectral signatures of these sparse pixels, we strictly excluded them due to the fundamental biophysical disparities between shrubs and trees—namely in carbon allocation, allometric scaling, and biomass turnover (Verbruggen et al., 2021). This distinction is critical given the vast distribution of shrublands in these regions; for instance, the data from the seventh NFI report ~7.87 Mha of shrubland versus only ~2.55 Mha of wooded land in Xinjiang, Qinghai, and Ningxia. By excluding

shrublands, we minimize commission errors and ensure the biophysical accuracy required for PFT mapping in DGVMs, thereby preventing systematic biases in simulated surface fluxes such as GPP and ET. Future iterations could mitigate this limitation by adopting a soft classification approach that quantifies sub-pixel fractional cover, providing a representation of vegetation transitions in sparse landscapes."

References:

Verbruggen, W., Verbeeck, H., Horion, S., Souverijns, N., and Schurgers, G.: Mapping Sahelian Ecosystem Vulnerability to Vegetation Collapse: Vegetation Model Optimization, *IEEE Int. Geosci. Remote Sens. Symp*., 1591–1593, https://doi.org/10.1109/IGARSS47720.2021.9554686, 2021.

Yu, Z., Ciais, P., Piao, S., Houghton, R. A., He, Y., Grimrim, H., Lewis, S. L., Li, Y., Yue, C., Wang, T., Li, W., Canny, C., Chen, A., Wang, S., Liu, L., Liu, Y., and Wang, Y.: Forest expansion dominates China's land carbon sink since 1980, *Nat. Commun.*, 13, 5374, https://doi.org/10.1038/s41467-022-32961-2, 2022.

How did you ensure that the definition from the technical standard (GB/T 38590-2020) aligns perfectly with the spectral signatures detectable by the satellite-based LULC products you used?

Response: We appreciate your insightful comments. Bridging the gap between ground-based inventory standards and optical satellite spectral signatures remains a fundamental challenge in remote sensing. Rather than assuming inherent consistency between the two, we developed a method based on the principle of "Statistics constrain quantity, remote sensing locates space." We actively bridge this gap through the following three strategies:

1. Quantity anchoring: We do not rely on the raw spectral classification of satellite products to determine absolute forest area, given the uncertainties arising from spectral confusion. Instead, we treat provincial NFI statistics as the absolute "ground truth." Regardless of the raw area derived from satellite inversions, our algorithm forces the final reconstructed forest area to strictly match the "wooded land" area recorded in the NFI. Consequently, our area standard is intrinsically governed by the NFI rather than raw satellite classifications.

2. Spatial Allocation: With the total quantity fixed, we utilize multisource satellite LULC products to determine the probability of a pixel being classified as forest (defined as classification consistency),

rather than accepting individual classification results. By calculating the "consistency" across multiple datasets, we identify areas with the highest probability of forest occurrence. Furthermore, we incorporate Growing-Season NDVI ranking as a secondary screening criterion, prioritizing pixels exhibiting the strongest forest spectral signatures (high NDVI). By integrating these spatial indicators to fulfill the NFI area quota, our approach leverages the spatial heterogeneity of remote sensing data within the rigid framework of statistical constraints. This ensures that the selected pixels represent areas with the highest vegetation vigor and biophysical consistency with "forests." Additionally, since the tree cover definitions of input satellite products (generally 15%–30%) largely envelope the NFI standard (>20%), our top-down filling logic naturally prioritizes core forest pixels that satisfy stricter definitions (high consistency), ensuring the rationality of spatial allocation.

3. Hierarchical constraints: The NFI definitions establish a rigorous hierarchical classification system for our study, clearly delineating the inclusion relationships among different forest types. These explicit hierarchical relationships provide constraints for the reconstruction of our forest and PFTs (forest → needleleaf/broadleaf → evergreen/deciduous), ensures the scientific robustness of long-term time series reconstruction. This approach achieves semantic alignment between NFI statistics and the "forest" (i.e., tree-covered areas) typically identified in satellite LULC products, effectively eliminating discrepancies caused by ambiguous shrub definitions in certain satellite datasets.

In summary, although raw definitions vary across satellite products, our "Ensemble Consistency + NFI Constraint" method forces the final map to strictly align with the GB/T 38590-2020 standard in terms of area, thereby minimizing the discrepancy between statistical definitions and spectral observations.

Response to comments

Paper: essd-2025-475

Title: 1 km annual forest cover and plant functional type dataset for China from 1981 to 2023

Journal: Earth System Science Data

Response to the Reviewers #2

RC2: Comment on essd-2025-475

This study presents a 1 km resolution annual forest cover dataset (1981–2023) and an associated plant functional type (PFT) dataset by integrating multi-source remote sensing data with National Forest Inventory (NFI) statistics. I have reviewed the authors' responses to the previous comments point by point, and most of the comments have been adequately addressed. The manuscript would benefit from minor language polishing, as some sentences are relatively long and could be made more concise to improve readability.

Response: We sincerely appreciate your acknowledgement of our previous revisions. We fully accept your suggestion regarding language quality. In this revision, we have carefully proofread the entire manuscript, specifically focusing on simplifying overly long sentences and editing for correct English language, grammar, punctuation, and phrasing to enhance conciseness and overall readability. All modifications are reflected in the tracked changes version of the manuscript.

[Specific comments 1] Line 11: The spatial resolution of the new forest dataset is 1 km, rather than approximately 1 km.

Response: We appreciate your suggestion. We have standardized the description of the spatial resolution to "1 km" in the manuscript. The relevant modification is located at Line 12:

"Here, we developed a high-resolution (1 km), annual forest cover dataset for China for 1981–2023."

[Specific comments 2] Lines 23–24: The statement describing the reduction of simulation errors for ET, LAI, and carbon flux currently reports a range of values. It would be clearer to provide the specific values for each variable, rather than only a range.

Response: We appreciate your suggestion. We agree that listing specific values for each variable significantly improves clarity compared to providing a range. We have revised this section to explicitly detail the percentages of error reduction for GPP, NEE, LAI, and ET. The relevant modifications are located at Lines 23–26:

"Specifically, it reduced errors relative to over extensive regions, outperforming these baselines across 77.7% and 85.2% of the terrestrial area for gross primary productivity (GPP), 63.1% and 69.7% for net ecosystem exchange (NEE), 66.9% and 77.3% for the leaf area index (LAI), and 78.7% and 85.3% for actual evapotranspiration (ET)."

[Specific comments 3] Line 180: The definition of "forest" or "wooded land" could be described in the manuscript rather than in supplementary materials.

Response: We appreciate your suggestion. We agree that providing clear definitions of "forest" and "wooded land" directly in the main text is essential for the clarity of the manuscript. We have moved the definition table (formerly Table S2 in the supplementary material) to the main text, where it is now designated as **Table 2**. Consequently, the number of subsequent tables in the main text and supplementary materials has been adjusted to ensure consistency.