General comments

This study presents a novel global 30 m resolution annual cropland dynamics dataset (GACED30) spanning 2000–2024, aiming to address the limitations of existing global cropland products in terms of temporal continuity, definitional consistency, and the monitoring of structural evolution. By constructing a gap-free SDC30 time series, introducing a spectral-semantic sample alignment strategy, and employing a CatBoost classifier combined with spatiotemporal post-processing, the authors generated the first annual, high-resolution cropland dynamics product with long-term continuity. The dataset demonstrates excellent performance in accuracy assessment and shows strong agreement with FAO national statistics, successfully revealing a “Global North-South divergence” pattern in cropland expansion and contraction. Overall, the methodology is innovative, the data product holds significant scientific value and application potential. Suggested to be published after revisions. The following are the questions and some mistakes in this manuscript.

> We sincerely thank your for the careful assessment of GACED30 and for the constructive recommendation. In response to these comments, we have revised both the manuscript and Supplement to improve methodological transparency and reproducibility, clarify data availability and the temporal coverage of the compared products, report the sample-screening, CatBoost, and CleanLab settings, provide an explicit operational interpretation and sensitivity analysis for the minimum slope threshold, address residual temporal mismatch in inter-product comparisons, strengthen the result-based discussion, improve the cartographic elements, standardize dataset and software references, and revise the English expression throughout. We have also moderated several overly strong statements and now distinguish more carefully among static thematic agreement, transition-detection performance, national-scale statistical consistency, and the remaining limitations of the dataset. These revisions have made the presentation more precise, reproducible, and scientifically bounded.

Comments

(1) Consider adding a brief statement on data availability at the end of the abstract for completeness.

> Thank you for this helpful suggestion. Although a brief repository link was already present in the final sentence of the Abstract, we agree that the identity of the data product and the location of its archived record should be stated as explicitly as possible. We therefore revised this sentence to name GACED30 directly and provide its persistent Zenodo DOI.

The manuscript has been revised as follows:

A. Abstract, Lines 13–29 – We revised the final sentence to identify GACED30 explicitly and provide its persistent repository link.

"GACED30 provides a harmonized baseline for monitoring global cropland-extent dynamics and is publicly available at https://doi.org/10.5281/zenodo.18199675 (Chen et al., 2026)."

(2) Table 1 is comprehensive, but consider adding a column indicating the temporal coverage of each product to facilitate direct comparison with GACED30's 25-year record.

> Thank you for this helpful suggestion. Temporal coverage was already included in the original "Temporal attributes" column. To make this information more explicit and facilitate direct comparison across products, we renamed the column as "Temporal frequency / coverage" and standardized its entries to report both the update frequency and coverage period."

(3) The preliminary CatBoost model used for sample screening lacks detailed parameter specifications. Providing these details or including them in supplementary materials would enhance reproducibility. The CleanLab-based temporal cleaning process is mentioned but not quantitatively described. What confidence thresholds were used? How were temporal inconsistencies identified? These should be clarified.

> Thank you for this important comment. We agree that the original manuscript did not provide sufficient implementation details for either the preliminary

CatBoost-based agreement screening or the subsequent CleanLab screening. We also agree that the original expression “temporal cleaning” could incorrectly imply that CleanLab directly evaluates transitions or imposes a temporal-consistency rule between adjacent years. We have therefore reorganized the description into a four-stage sample-construction procedure, reported the quantitative settings in Sect. S1.1 and Table S1, and replaced the ambiguous terminology with “year-specific weak-label screening.”

The preliminary agreement-screening classifier, the CatBoost models used to generate out-of-fold probabilities for CleanLab, and the final annual classifiers shared the following core configuration: 2,000 iterations, a tree depth of 8, a learning rate of 0.05, an L2 leaf-regularization coefficient of 6.0, Logloss as the loss function, F1 as the evaluation metric, a random seed of 42, 80 CPU threads, and CPU execution. The role of the preliminary classifier is now identified explicitly as Stage 1: it was trained exclusively with the expert-annotated FAST-Crop samples and retained a candidate only when its predicted Cropland/Non-cropland class agreed with the corresponding GLAD-stratified label.

We further added a detailed quantitative description of Stage 2, which performs spatial, spectral, and category filtering after the preliminary agreement screening. The land surface was divided into 5°×5° geographic grids. Within each grid, the six SDC30 surface-reflectance bands were transformed as $z = \log(1 + SR)$, and FAST-Crop Cropland samples were used to estimate the corresponding multivariate Gaussian reference distribution. An augmented Cropland candidate passed the hard spectral prefilter only when its squared Mahalanobis distance satisfied $D^2 \leq \chi^2(0.999, 6) = 22.457744$. Candidates were rejected when no reliable exact-grid, parent-grid, or global reference distribution was available, when D^2 exceeded this threshold, or when the log-likelihood was non-finite. Passing candidates were sampled according to their Gaussian likelihood using a likelihood temperature of 1, after which category balancing was conducted within each grid according to the Cropland/Non-cropland proportions represented by FAST-Crop.

For Stage 4, each retained augmented location was assigned annual weak labels for 2000–2024 using the relevant GLAD Cropland epoch. For each year, five-fold out-of-fold CatBoost probabilities were generated using StratifiedKFold(n_splits=5, shuffle=False) and supplied to CleanLab's default find_label_issues procedure. FAST-Crop labels were treated as fixed reference labels and were excluded from pruning, whereas only augmented point-year labels were eligible for screening. No manually selected probability or confidence threshold was applied. A flagged augmented point-year was excluded from the training sample set of the corresponding annual classifier, and an augmented location was removed entirely when more than 25% of its 25 annual labels were flagged, corresponding to at least seven years. After this procedure, the final augmented sample set contained 332,471 locations, including approximately 25,000 Cropland and 307,000 Non-cropland samples. CleanLab therefore identifies year-specific weak-label conflicts, in which the assigned GLAD-derived label is not supported by the spectral evidence for that year; it does not identify a temporal inconsistency through an explicit year-to-year transition rule. Direct temporal information is introduced separately during post-classification refinement through the limited rotation rule in Sect. 3.3.1. We have revised the manuscript accordingly to distinguish sample screening from temporal refinement.

The manuscript has been revised as follows:

A. Sect. 3.2, Lines 269–273 – We replaced the previous "temporal refinement" description with year-specific weak-label screening and clarified that direct temporal linkage is introduced only during post-classification refinement.

"To generate annual cropland probability maps under a common mapping framework, we first constructed spatially extensive and thematically harmonized annual training subsets through multi-stage sample alignment and year-specific weak-label screening (Sect. 3.2.1). We subsequently used year-specific CatBoost models to estimate pixel-wise cropland probabilities across

the 2000–2024 archive (Sect. 3.2.2). The annual classifiers were trained independently, and direct temporal linkage between their outputs was introduced only through the limited post-classification rotation rule described in Sect. 3.3.1."

B. Sect. 3.2.1, Lines 285–294 – We added the main-text description of Stage 2 spatial, spectral, and category filtering and Stage 3 performance-guided sample optimization, with detailed settings directed to Sect. S1.1 and Table S1.

"In Stage 2, spatial, spectral, and category filtering was performed. The global land surface was divided into 5° × 5° geographical grids. Within each grid, augmented candidates were sampled according to the statistical distribution of the six-band annual-median SDC30 surface-reflectance vector, drawing on the approximately Gaussian behaviour of log-transformed surface reflectance in homogeneous land-cover regions (Liao et al., 2026a). Category balancing was then performed by sampling the filtered augmented candidates according to the Cropland/Non-cropland proportions of FAST-Crop within each grid. In Stage 3, the remaining candidates were optimized using the sample-size–performance scaling relationship established by Liao et al. (2026b). At each iteration, alternative 10,000-sample batches were evaluated, and the batch producing the highest estimated asymptotic performance was retained. Iterations were stopped once the retained augmented-sample total reached the target of approximately 340,000. Detailed implementation settings are provided in Sect. S1.1 and Table S1."

C. Sect. 3.2.1, Lines 301–314 – We added the annual weak-label assignment, five-fold out-of-fold prediction design, CleanLab operation, pruning scope, point-year exclusion rule, and location-level removal criterion.

"Finally, annual weak labels were assigned directly to the selected spatial samples for 2000–2024. For 2000–2019, the label from each GLAD Cropland epoch was assigned to every year covered by that epoch; because GLAD

Cropland provides no subsequent epoch, the 2016–2019 labels were also used as the initial weak labels for 2020–2024. For each year, five-fold out-of-fold CatBoost probabilities were generated using StratifiedKFold(n_splits=5, shuffle=False) and supplied to CleanLab's default find_label_issues procedure (Northcutt et al., 2021). FAST-Crop labels were treated as fixed reference labels and were excluded from possible pruning, whereas only the augmented point-year labels were eligible for screening. No manually selected probability or confidence threshold was applied. An augmented point-year flagged by CleanLab was excluded from the training sample set for the corresponding annual classifier, and an augmented location was removed entirely when more than 25% of its 25 annual labels were flagged, corresponding to at least seven years. This procedure screens year-specific weak-label conflicts rather than imposing a transition rule between adjacent years. The annual training subsets therefore shared the same expert-reference basis, feature definition, sample-construction framework, and screening procedure, but were not identical because augmented point-year labels could be excluded separately for each annual classifier. After annual weak-label assignment and screening, the final augmented sample set contained 332,471 locations, comprising approximately 25,000 Cropland and 307,000 Non-cropland samples."

D. Sect. 3.2.2, Lines 316–320 – We clarified that the preliminary screening classifier, the out-of-fold CatBoost models used for CleanLab, and the final annual classifiers shared the same reported core configuration.

"Following construction of the annual training subsets, we trained independent year-specific classifiers to account for interannual phenological variability. For each year from 2000 to 2024, the corresponding valid FAST-Crop and augmented samples were used to train a CatBoost classifier (Prokhorenkova et al., 2018). The preliminary screening classifier, the models used to generate out-of-fold probabilities for CleanLab, and the final annual classifiers shared the same core CatBoost configuration, which is reported in Sect. S1.1 and Table

S1.”

E. Sect. 3.3.1, Lines 344–352 – We clarified that temporal information is incorporated after annual classification through a limited rotation rule rather than through CleanLab.

“Because the annual CatBoost models were trained independently, temporal evidence was introduced after classification through a limited rotation rule targeting uncertain observations. Specifically, we identified pixels within the probability interval ($P \in [0.3, 0.5]$), which would otherwise be classified as Non-cropland under the standard binary threshold.

For a pixel in year t, an uncertain observation within this interval was retained as Cropland under a temporary-fallow interpretation only when the same pixel was classified as active cropland ($P>0.5$) in both the preceding year t-1 and succeeding year t+1. This rule bridges isolated one-year gaps that are compatible with temporary fallow or weak phenological signals. The temporal context reduces, but does not eliminate, confusion with spectrally similar pasture, rangeland, or bare surfaces. The rule does not alter observations with probabilities below 0.3, smooth complete pixel trajectories, or impose general temporal consistency across the annual series.”

F. Sect. S1.1, Lines 15–59 – We added the detailed four-stage construction procedure. In particular, Stage 1 specifies the preliminary CatBoost agreement screening, Stage 2 reports the 5°×5° spatial stratification, six-band transformation, Gaussian reference distribution, Mahalanobis-distance threshold, likelihood sampling, and category-balancing rules, and Stage 4 specifies the annual CleanLab screening.

G. Table S1, Lines 60 – We added the complete CatBoost configuration together with the principal Stage 2 filtering, sample-optimization, and CleanLab settings.

(4) The selection of β = 0.2 ha/year as the threshold for significant change (within a 1 km grid) lacks explicit justification. Please provide reasoning for this threshold (e.g., based on expected noise levels or agronomic considerations).

> Thank you for identifying that the original justification of the threshold was insufficient. We agree that the previous statement—that 0.2 ha yr^{-1} was selected to filter out fluctuations caused by temporary fallowing or inter-annual classification instability—was not adequately supported by an explicit noise calibration or an agronomic benchmark. We have therefore removed this unsupported claim and substantially revised Sect. 3.4.

First, we now distinguish the estimated Theil–Sen slope β from the selected minimum effect-size threshold τ. Statistical significance is determined by the Mann–Kendall p-value, whereas τ specifies the minimum trend magnitude interpreted as grid-scale structural change. We also distinguish the general structural trend, defined by the sign and magnitude of β, from the stringent structural trend, which additionally requires $p<0.01$.

We further clarify that neither the Mann–Kendall test nor the Theil–Sen estimator provides a theoretically unique or universally optimal value of τ. In this study, τ=0.2 ha yr^{-1} was adopted as an operational threshold with an interpretable grid-scale meaning. A 1×1 km grid covers approximately 100 ha, so this value represents a linear trend equivalent to 0.2% of the grid area per year and 4.8 ha, or 4.8% of the grid area, across the 24 annual intervals between 2000 and 2024. Although 0.2 ha is numerically equivalent to approximately 2.2 nominal 30 m pixel areas, τ is estimated from the complete 25-year trajectory rather than applied as a single-year pixel-count threshold.

To make the dependence on this operational choice transparent, we added sensitivity analyses for τ=0.10–0.50 ha yr^{-1} at intervals of 0.05 ha yr^{-1}. Increasing τ progressively reassigns lower-magnitude expansion and reduction grid cells to stable, as expected. Nevertheless, the continental composition of the detected changes varies only modestly, and nearby thresholds retain high spatial overlap with the τ=0.2 reference masks. At τ=0.15 and 0.25, the Jaccard indices are approximately 0.88–0.91 for the general masks and 0.91–0.93 for the stringent masks. These results support $\tau=0.2$ as a reasonable and interpretable operational reference for this analysis, without implying that it is uniquely optimal or universally applicable.

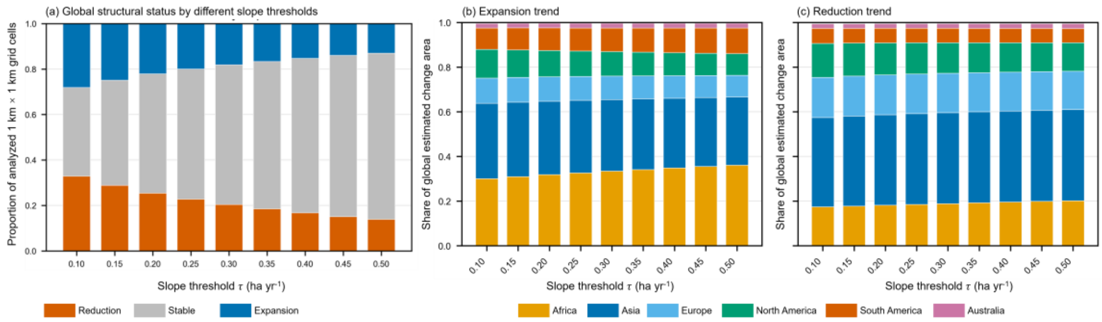

Figure S1. Sensitivity of global structural evolution trend composition and the continental distribution of cropland change to the minimum slope threshold. (a) Proportions of analysed 1 × 1 km grid cells assigned to general expansion ($\beta>\tau$), general reduction ($\beta<-\tau$), and general stable ($|\beta|\leq\tau$) as τ varies from 0.10 to 0.50 ha yr−1. This panel applies the general structural trend definition independently of MK significance. (b–c) Continental shares of the total global area classified as expansion and reduction, respectively, under the corresponding thresholds. For each threshold, the continental shares within each change class sum to 1. Panels (b–c) describe the continental composition of the globally detected change area.

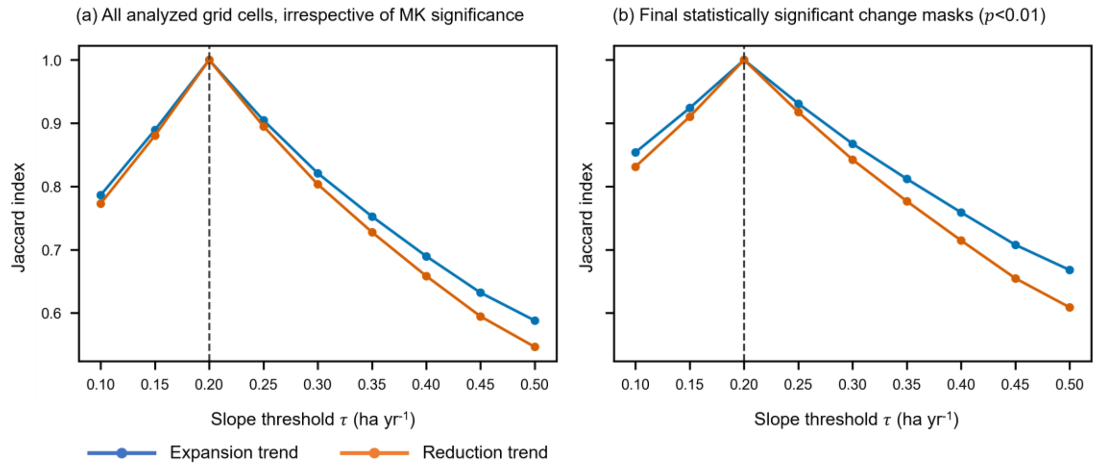

Figure S2. Sensitivity of the spatial expansion and reduction masks to the minimum slope threshold. The conventional Jaccard index, $J=|M_\tau \cap M_{0.2}|/|M_\tau \cup M_{0.2}|$, compares the mask obtained at each τ with the corresponding reference mask at $\tau=0.2$. Results are shown for (a) all analysed 1 × 1 km grid cells classified according to the sign and magnitude of the Theil–Sen slope, irrespective of Mann–Kendall significance, and (b) the final statistically significant change masks, for which expansion requires $p<0.01$ and $\beta>\tau$, and reduction requires $p<0.01$ and $\beta<-\tau$. A Jaccard index of 1 indicates *identical masks.*

The manuscript has been revised as follows:

A. Sect. 3.4, Lines 373–394 – We replaced the original threshold passage and revised the structural-trend definition.

"To resolve this ambiguity and identify the directional evolution of these landscapes, we derived Structural Evolution Indicators from GACED30 using a grid-based trend analysis framework (Fig. 2d). Area calculations were performed in a global equal-area projection following the practical principles outlined by Tyukavina et al. (2025). The annual 30 m binary cropland maps were aggregated to 1 × 1 km grid cells, and mapped cropland area was summed within each grid cell for each year. This produced a 25-observation annual cropland-area series for each analysed grid cell over 2000–2024. The non-parametric Mann–Kendall (MK) test and Theil–Sen estimator were applied to each time series, yielding the MK significance level (p), the Theil–Sen slope (β, ha yr^{-1}), and the confidence interval of the slope. Using the minimum slope threshold τ, grid cells with $\beta>\tau$, $\beta<-\tau$, and $|\beta|\leqslant\tau$ were classified as general expansion, general reduction, and stable, respectively. Expansion and reduction were additionally classified as stringent when the corresponding Mann–Kendall test satisfied $p<0.01$.

Neither the MK test nor the Theil–Sen estimator provides a theoretically unique or universally optimal value of τ. The appropriate minimum effect size depends on the spatial aggregation unit, study duration, and analytical objective. For the principal analysis in this study, we adopted $\tau=0.2$ ha yr^{-1} as an operational grid-scale threshold. Because a 1 × 1 km grid cell covers approximately 100 ha, this value corresponds to a linear trend equivalent to 0.2% of the grid area per year and to 4.8 ha, or 4.8% of the grid area, across the 24 elapsed annual intervals between 2000 and 2024. It therefore provides an interpretable minimum magnitude for identifying landscape-scale structural change over the study period.

Although 0.2 ha is numerically equivalent to approximately 2.2 nominal 30 *m pixel areas, τ is not applied as a single-year pixel-count detection limit. Instead, β is estimated from the complete 25-year cropland-area trajectory within each grid cell. We do not consider τ=0.2 ha yr^{-1} to be universally optimal, and other thresholds may be appropriate for analyses with different objectives or tolerances for lower-magnitude change. We therefore evaluated the sensitivity of both the general and stringent structural-evolution classifications by varying (τ) from 0.10 to 0.50 ha yr^{-1} at intervals of 0.05 ha yr-1 (Figs. S1 and S2).”*

B. Sect. 4.3.3, Lines 767–774 – We revised the opening results paragraph to report both the general and stringent structural-trend definitions and the class composition under τ=0.2 ha yr^{-1}.

“Leveraging the continuous annual time series of GACED30, we quantified global cropland dynamics through grid-level MK and Theil–Sen trend analysis rather than simple bi-temporal differencing. Fig. 8a presents both complementary interpretations of the resulting trends: the general structural trend describes all grid cells whose estimated slope exceeds the selected minimum magnitude, whereas the stringent structural trend identifies the subset that additionally satisfies $p<0.01$. Under the principal threshold of τ=0.2 ha yr^{-1}, stable was the dominant structural-status class, accounting for slightly more than half of the analyzed grid cells; approximately one-quarter were classified as general reduction and slightly more than one-fifth as general expansion (Fig. S1a). Stable cropland was especially extensive in established agricultural regions such as the US Corn Belt, the North China Plain, and Western Europe.”

C. Sect. 4.3.3, Lines 775–786 – We added a separate paragraph reporting the threshold-sensitivity results, including the progressive reassignment of change cells to stable, continental composition, and Jaccard overlap with the τ=0.2 masks.

“Because τ is an operational rather than theoretically optimal threshold, we

examined whether the principal geographical interpretation depended strongly on its selected value. Increasing τ from 0.10 to 0.50 ha yr^{-1} progressively reassigned lower-magnitude expansion and reduction grid cells to stable, as expected, and therefore changed the absolute mapped extent of structural change (Fig. S1a). However, the continental composition of the general expansion and reduction areas changed only modestly across the evaluated thresholds, with no abrupt redistribution around τ=0.2 ha yr^{-1} (Fig. S1b–c). Spatial agreement with the τ=0.2 reference masks was also high for nearby thresholds under both definitions: at τ=0.15 and 0.25, the Jaccard indices were approximately 0.88–0.91 for the general masks and 0.91–0.93 for the stringent masks (Fig. S2). Thus, the precise extent of lower-magnitude change remains threshold-dependent, but the broad continental distribution and principal spatial patterns are retained across reasonable values surrounding τ=0.2. Together with the regional patterns in Fig. 8, this supports the robustness of the central interpretation that agricultural expansion is disproportionately concentrated in the Global South, whereas widespread stability and localized contraction characterize much of the Global North. Nevertheless, the sensitivity results do not imply that τ=0.2 is a uniquely optimal threshold."

D. Sect. 3.4, Lines 395–399 – We revised the methodological qualification of the complete annual trajectory to avoid claiming that trend analysis guarantees a permanent conversion or independently determines its cause.

"Compared with bi-temporal differences between multi-year epochs, analysis of all 25 annual observations reduces sensitivity to the selection of baseline and endpoint windows and limits the influence of short-lived fluctuations. The Mann–Kendall test and Theil–Sen estimator quantify the statistical support, direction, and magnitude of mapped trends under the specified criteria. These statistics do not independently establish the cause of a trend or guarantee that every mapped change represents a permanent land-cover conversion."

E. Sect. 5.1, Lines 813–819 – We revised the Discussion to state that the full trajectory reduces sensitivity to endpoint selection while the exact spatial extent of expansion and reduction remains threshold-dependent.

"The complete annual record supports two complementary descriptions of cropland dynamics. Cropping Frequency summarizes the proportion of years in which a pixel was mapped as Cropland, whereas the Structural Evolution Indicators use annual grid-level cropland area, the Mann–Kendall test, and Theil–Sen slopes to identify persistent directional patterns. Analysis of the full trajectory reduces sensitivity to the selection of individual endpoint years, but the exact spatial extent of expansion and reduction remains dependent on the minimum slope threshold. The sensitivity analysis indicates that the broad continental distribution is retained across thresholds near τ=0.2 ha yr^{-1}, without implying that this value is uniquely optimal or that the detected trends establish their underlying causes."

(5) The comparison with other products (Table 2) is valuable, but some products (e.g., GLAD Cropland, ESA World Cover) have differing temporal baselines. The potential impact of temporal mismatches on comparative metrics should be explicitly addressed.

> Thank you for highlighting this important limitation. We agree that differing temporal baselines mean that the comparison in Table 2 is not perfectly contemporaneous and that nearest-date matching can reduce, but cannot eliminate, the resulting temporal offset. We also acknowledge that the original wording could be read as overstating the relative accuracy of GACED30. We have therefore removed language implying statistically significant or universal superiority over all peer products. The revised manuscript interprets Table 2 as evidence of relative static thematic agreement with the adopted FAST-Crop/GACED30 reference definition, rather than as proof that GACED30 is universally more accurate across all cropland categories, years, or change conditions. This limitation is stated directly in the interpretation of Table 2 in Sect. 4.1.1 and is carried into the broader discussion of incomplete temporal

comparability and reference coverage in Sect. 5.3.

The FAST-Crop validation samples predominantly represent land-cover conditions around 2015. For each product, we therefore selected the available annual layer or multi-year epoch closest to the reference-imagery year and constructed one confusion matrix for that product. This procedure minimizes temporal offset, but genuine cropland changes occurring between the reference date and the selected product layer or epoch can still be recorded as disagreement. Consequently, part of the difference among comparative metrics may reflect residual temporal mismatch in addition to mapping behaviour and cropland-definition differences.

The manuscript has been revised as follows:

A. Sect. 3.5.1, Lines 409–413 – We clarified the nearest-date matching strategy used to compare products with different temporal availability.

"We evaluated the pixel-level accuracy of GACED30 using the independent FAST-Crop validation set. To ensure a consistent evaluation across different satellite products with varying availability, we employed a nearest neighbor temporal matching strategy. For any given validation sample labeled in year T_sample, we selected the map layer from the product (GACED30 or reference products like GLAD Cropland) with the year T_product that minimized the absolute temporal difference |T_sample-T_product |. This ensures that the validation reflects the product's performance closest to the ground reference date."

B. Sect. 4.1.1, Lines 552–558 – We explicitly acknowledged that temporal and definitional mismatches can contribute to disagreement and qualified the interpretation of Table 2 so that it is not presented as evidence of universal superiority.

"Part of the mismatch with peer products may reflect differences in cropland definitions rather than mapping errors alone. Specifically, GACED30 includes orchards and active fallow but excludes temporary meadows and pastures,

whereas peer products differ in their treatment of these categories. Temporal mismatches between the validation samples and product layers may also contribute to the observed differences. The metrics in Table 2 should therefore be interpreted as the products' relative agreement with the adopted FAST-Crop/GACED30 definition rather than as evidence of universal superiority across all cropland categories. Accordingly, Table 2 supports the static thematic reliability of GACED30 under the adopted Cropland/Non-cropland reference definition but does not directly assess its ability to detect inter-period cropland changes."

C. Sect. 5.3, Lines 868–875 – We acknowledged the broader limitation that complete interannual comparability is not enforced and that available multi-temporal references do not validate every annual layer or exact transition year. *"Temporal comparability is supported by the common observation, feature, sample-construction, and model protocols, but complete pixel-level interannual consistency is not enforced. The independently trained annual classifiers can remain sensitive to annual spectral conditions, particularly near the classification threshold or during abrupt land-cover changes. Differences between the earlier Landsat–MODIS observation environment and the more recent sensor era may also influence temporal comparability and are not eliminated solely by trend analysis. Moreover, the available GLAD Cropland and LUCAS references assess crop gain and crop loss across multi-year intervals rather than every annual layer. They cannot identify the exact transition year, detect all within-interval reversals, or distinguish permanent cropland reduction from temporary non-cropland states. More independent annual and transition-specific reference observations are therefore needed."*

(6) Please further discuss based on the results of this study to enhance the depth of the discussion.

> Thank you for this helpful suggestion. We agree that the original Discussion focused primarily on the general implications of GACED30 and did not

sufficiently connect those implications to the principal results of this study. We have therefore rewritten Sect. 5.2 around the modest increase in global cropland area, the predominance of stable cropland, and the spatial concentration of expansion and reduction. The revised discussion explains how the coexistence of relative global stability and substantial regional reorganisation can inform food-security assessment, agricultural-frontier monitoring, land-use policy evaluation, carbon accounting, and Earth system model benchmarking. We also state the complementary data required for these applications and avoid interpreting temporal correspondence as evidence of causality.

The manuscript has been revised as follows:

A. Sect. 5.2, Lines 833–847 – We connected the observed modest global net increase and concentrated regional reorganisation to food-security assessment, agricultural-frontier monitoring, and land-use policy analysis, while defining the limits of those applications.

"The results in Sects. 4.3.1–4.3.3 indicate that global cropland increased modestly from 1458.5 Mha in 2000 to 1488.5 Mha in 2024, while stable cropland remained predominant and statistically significant expansion and reduction were spatially concentrated. This combination shows that a relatively small global net change can coexist with substantial regional reorganisation, particularly the contrast between concentrated expansion in parts of Africa and South America and widespread stability or reduction in much of the Global North. Because GACED30 maps cropland extent rather than crop type, yield, production, or food access, it cannot measure food security directly. When combined with crop-type maps, yield and production statistics, population, market-access, and nutrition data, however, it can provide the land-footprint component of food-security assessments and help distinguish cases in which production changes are accompanied by cropland expansion from those occurring without marked extent change. The significant trend layer can also identify candidate agricultural-frontier and persistent cropland-reduction zones

for targeted monitoring; determining the preceding land cover and environmental implications requires independent forest, grassland, biodiversity, protected-area, infrastructure, and land-tenure information. When combined with trade-flow and socioeconomic data, these spatial trajectories can support telecoupled land-system analyses linking distant production and consumption regions; when paired with policy boundaries, implementation dates, and an appropriate comparison or counterfactual design, they can also serve as an outcome layer for land-use policy evaluation, although GACED30 alone cannot establish causal effects."

B. Sect. 5.2, Lines 848–857 – We explained how the structural-trend record can support carbon-accounting sensitivity analyses and Earth system model benchmarking when combined with appropriate complementary datasets.

"The structural trend record also has practical value for carbon accounting and land-use–climate analysis because persistent directional changes can be distinguished from short-term fluctuations under a transparent statistical criterion. When combined with preceding land-cover classes, biomass and soil-carbon stock data, emission factors, and model assumptions, mapped cropland expansion or reduction can provide spatial activity information or stratification for carbon-accounting and sensitivity analyses; GACED30 itself does not quantify carbon stocks or emissions. At coarser scales, the 2000–2024 record can be aggregated and harmonised with model grids to compare the spatial allocation, direction, and persistence of observed cropland trends with land-use inputs or outputs used by Earth system models, including LUH2 (Hurtt et al., 2020). Such comparisons can support evaluation of land-use inputs and model behaviour, but GACED30 is not a complete land-use forcing dataset and cannot replace datasets that represent crop types, pasture, management, and transitions among non-cropland classes."

(7) The scale is missing in all the spatial maps. Some small maps lack the north

pointer or latitude/longitude.

> Thank you for this helpful comment. We carefully checked the spatial map panels throughout the manuscript. Scale bars and north pointers have been added to the relevant regional maps.

(8) Ensure consistency in formatting, particularly for datasets and software references (e.g., NASA JPL NASADEM citation).

> Thank you for this helpful suggestion. We have carefully checked and standardised the formatting of all references throughout the manuscript, including datasets and software-related references. In particular, the NASA JPL NASADEM citation has been revised to follow the recommended dataset citation format, with consistent author, dataset, repository, persistent identifier, and year information.

(9) Please improve English expression in the manuscript.

> Thank you for this helpful suggestion. We have carefully reviewed and revised the entire manuscript to improve English expression, grammatical accuracy, clarity, and overall fluency. We also standardized terminology and phrasing where appropriate.