Reviewer #2:

General Comments:

Comment #1

The manuscript describes a 30-m global forest disturbance dataset (11 disturbance types) for the time of 2000 to 2020. Disturbance is derived from Landsat data applying the CCDC analysis. My comments focus primarily on the accuracy assessment and area estimation components of the work. A primary area of improvement of the manuscript would be to provide a clear articulation of the sampling design used to collect the data for the accuracy assessment and area estimates. Without a clear description of the sampling design and additional details, it is impossible to ascertain how the accuracy and area estimates were obtained.

Response #1

Thanks for your constructive suggestions, which make our manuscript stronger. We have incorporated all of your suggestions. The primary objective of this study is to produce a global map of forest disturbance, which aligns closely with the title of our manuscript. Accordingly, in response to specific comment 4 (comment #10), we have removed the section on area estimation as it was not directly relevant to the main focus of the paper. Additionally, we have provided a detailed description of the sampling of sample points used for map production and validation. This addition enhances the clarity of the mapping methodology and improves the overall accuracy of the data. In this version, we have further revised the manuscript and addressed all your concerns. Please see the detailed point-by-point responses below.

Major Comments

Comment #1

Additional details related to the sampling design(s) must be provided. It is unclear how specifically the sample of 57,000 30-m sample units were selected for the model training and validation (Lines 72-73). In Section 2.3 (Lines 153-154), the text states that "8 individuals were uniquely responsible for selecting 8 types, while an additional 4 individuals conducted secondary confirmation of the selected samples." This text seems to be referring to the process of labeling the sample units, not explaining how the sample units were selected. Did these individuals actually choose which sample units (30-m pixels) were in the sample? There is no mention of randomization in the protocol for selecting the sample, and no details presented of whether strata are present, even though later in the manuscript stratified estimation formulas for accuracy metrics are provided (equations 5 through 10). To compound the confusion, the

Figure 3 confusion error matrix has a sample size of nearly 17,000, but there is no mention in the text of how these sample units were selected. Is it a random subset of the 57,000 mentioned earlier? Or are these 17,000 sample units entirely independent of the training sample of 57,000? It is essential to describe the sampling design(s) used to select these units. (**Revised**)

Response #1

Thanks for your constructive suggestions, which make our manuscript stronger. We have now included a detailed description of the sampling point selection process in the manuscript (Page 8-9, Line 173-182; Appendix A). The 57,000 sample points were selected through manual interpretation. Specifically, we first generated preliminary pre-labeled points automatically and globally with the support of existing single-type auxiliary datasets, such as forest fire map. These pre-labeled points, though not fully accurate, served as initial references. Our interpreters then visually identified and marked the final accurate sample points within the vicinity of these pre-labeled locations. This approach not only improves labeling efficiency but also ensures both the accuracy and randomness of the sample points.

"The selection of sample points was conducted in two steps: automated random generation and visual verification. We automatically generated imprecisely pre-labeled sample points through global simple random sampling, based on the existing auxiliary dataset of single-class disturbances, such forest fire, plantation, and cropland occupation, etc. Although these pre-labeled sample points are uncertainty, they provide our interpreters with rapid and spatially randomized regions for sample point selection. Subsequently, our interpreters performed visual verification and interpretation to mark accurate sample points within the vicinity of these pre-labeled locations. This approach significantly enhances both the efficiency and randomness of the sampling process. Undisturbed sample points are randomly generated and validated as stable pixels based on forest regions outside the Hansen's global forest change dataset. The sample point are evenly distributed in the global forest disturbance area (Appendix A).

For the challenging distinction of 'shifting cultivation', its identification relied on detecting unique cyclical patterns in the time series. Interpreters were trained to confirm three key characteristics within the high-resolution historical imagery. (1) Clear cyclical boundaries: evidence of alternating phases of forest (fallow), clearing/burning (clearance), and crops (cultivation) on the same parcel of land over multiple years; (2) Short-cycle land cover change: a complete cycle typically lasts a few years, distinguishing it from permanent deforestation for agriculture; and (3) Small-scale and fragmented spatial patterns: shifting cultivation plots are usually small, irregularly shaped, and interspersed with patches of mature forest. Sample points were only designated as shifting cultivation if they met multiple of these criteria *simultaneously to ensure accuracy. For the forest weak disturbance types caused by drought disturbance (16) and pest disturbance (17), their sample point selection needs to refer to high-resolution long-term remote sensing images."*

"Appendix A

The selection of sample points was primarily based on the time-series changes observed in Landsat images from 2000 to 2020, supplemented by historical high-resolution imagery from Google Earth. Through extensive analysis, eight types of forest disturbances were preliminarily identified: undisturbed (0), shifting cultivation disturbance (11), forestry replanting (12), plantation disturbance (13), deforestation of natural forests (14), forest fire disturbance (15), built-up area expansion (18), and cropland occupation (19). A total of 57,000 sample points representing these disturbance types were visually interpreted. These sample points are evenly distributed across global forest disturbance areas (Fig. A).

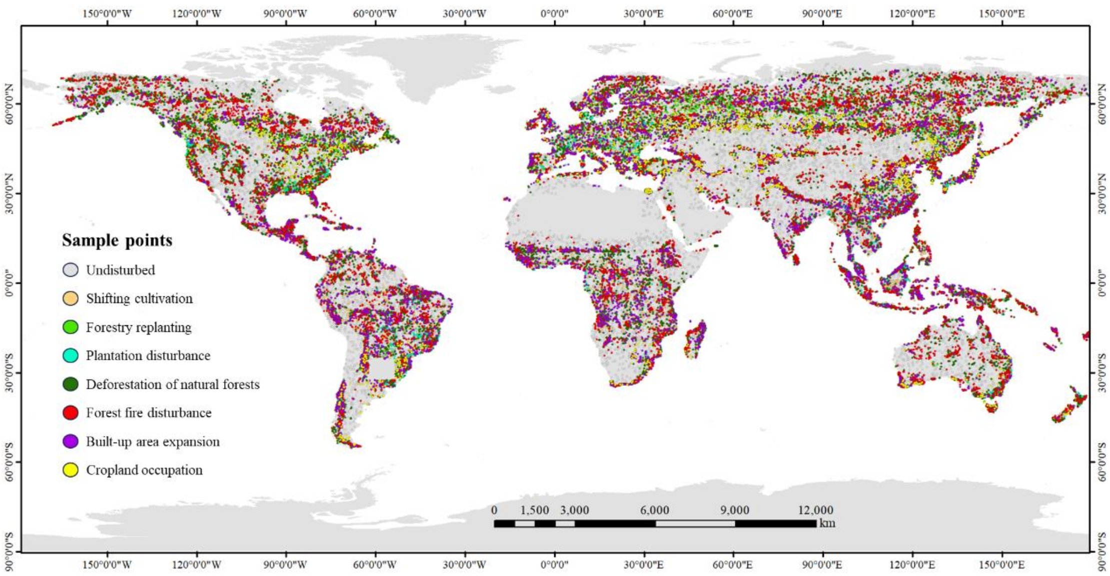


Figure A Overall spatial consistency comparison with CDGFL under logarithmic scale."

For the 17,000 validation sample points, the total set of 57,000 accurately interpreted samples in this study was randomly divided into two groups at a ratio of 7:3. A total of 40,000 samples were used for model training, and the remaining 17,000 were reserved for validation of the model results. This procedure was described in our previous version. We have further supplemented the relevant content in the manuscript (Page 10, Line 208-210).

"During the training and validation process of the model, the sample points we selected were divided into a training set and a validation set according to a 7:3 pattern. 40000 sample points

are used for model training, and another 17000 sample points are used to validate the model training results."

Comment #2

I have several concerns with the Figure 3 confusion matrix, which I will list as separate items as follows:

a) It seems very unlikely that there would be no errors associated with the undisturbed class (which is class 0). Out of 3476 cases, there was never a commission error or omission error of "undisturbed" – this class is perfectly mapped. It seems implausible that disturbed and undisturbed forest can be classified with 100% accuracy. (**Revised**)

Response #2

Thanks for your comment. This result is attributed to the application of masking procedures. This approach was designed to guarantee the reliability of the training samples, thereby improving the accurate identification of disturbed forest areas. We have supplemented relevant content in the manuscript (Page 12, Line 271-273; Page 9, Line 180-181; Page 10, Line 218-219). To ensure sample accuracy, the visually interpreted samples for the undisturbed category were deliberately selected from pixels exhibiting long-term stability. These pixels were also subjected to masking. We utilized forest change datasets, such as the Hansen et al. dataset, to exclude pixels that had undergone changes. Furthermore, the same masking protocol was applied to our final mapping products. Consequently, it is expected that the undisturbed class demonstrates nearly 100% accuracy.

"Here, we restricted our analysis to areas that had undergone forest disturbance. The observed 100% accuracy for the undisturbed class is attributable to the masking procedure applied using Hansen's Global Forest Change dataset."

"Undisturbed sample points are randomly generated and validated as stable pixels based on forest regions outside the Hansen's global forest change dataset. "

"The final GFD map, except for the newly added forest, uses Hansen's global forest change dataset as a mask to remove pixels that have not undergone forest changes."

Comment #3

b) The confusion matrix is presented in terms of sample counts, which is reasonable if the sampling design is simple random. Yet the authors present formulas for stratified sampling (equations 5-10). In particular, equation (5) indicates how the cell proportions should be estimated for a stratified sample, but that formula was not apparently used in the analysis. The confusion matrix should be presented in terms of the estimated pij (cell proportions) when stratified sampling is used. This concern links to comment 1 because the manuscript does not include description of the sampling design. (**Revised**)

Response #3

Thanks for your suggestion. As stated in the response to comment 1, we adopted a simple random sampling design. We appreciate your suggestion, which will greatly help improve our manuscript. We have modified the corresponding formula (Page 10-11, Line 230-244).

Overall accuracy (OA) derived from the overall error matrix of 11 forest disturbance types:

$$OA = \frac{\sum_{j=1}^{11} n_{jj}}{N} \tag{2}$$

User's accuracy for Class i (U_i) (the number of sample points mapped to class i with reference to class i)

$$U_i = \frac{n_{ii}}{n_{i.}} \tag{3}$$

Producer's accuracy for class j (the number of sample points with class j mapped to reference class j)

$$P_j = \frac{n_{jj}}{n_{.j}} \tag{4}$$

2.5.2 Estimating accuracy

The sampling variability associated with the accuracy estimates should be quantified by reporting standard errors. The variance estimators are provided below, and taking the square root of the estimated variance results in the standard error of the estimator. For overall accuracy, the estimated variance is:

$$V(O) = \frac{OA(1 - OA)}{N - 1} \tag{5}$$

For user's accuracy of map class i, the estimated variance is

$$V(U_i) = \frac{U_i(1 - U_i)}{n_{i.} - 1} \tag{6}$$

For producer's accuracy of reference class j, the estimated variance is

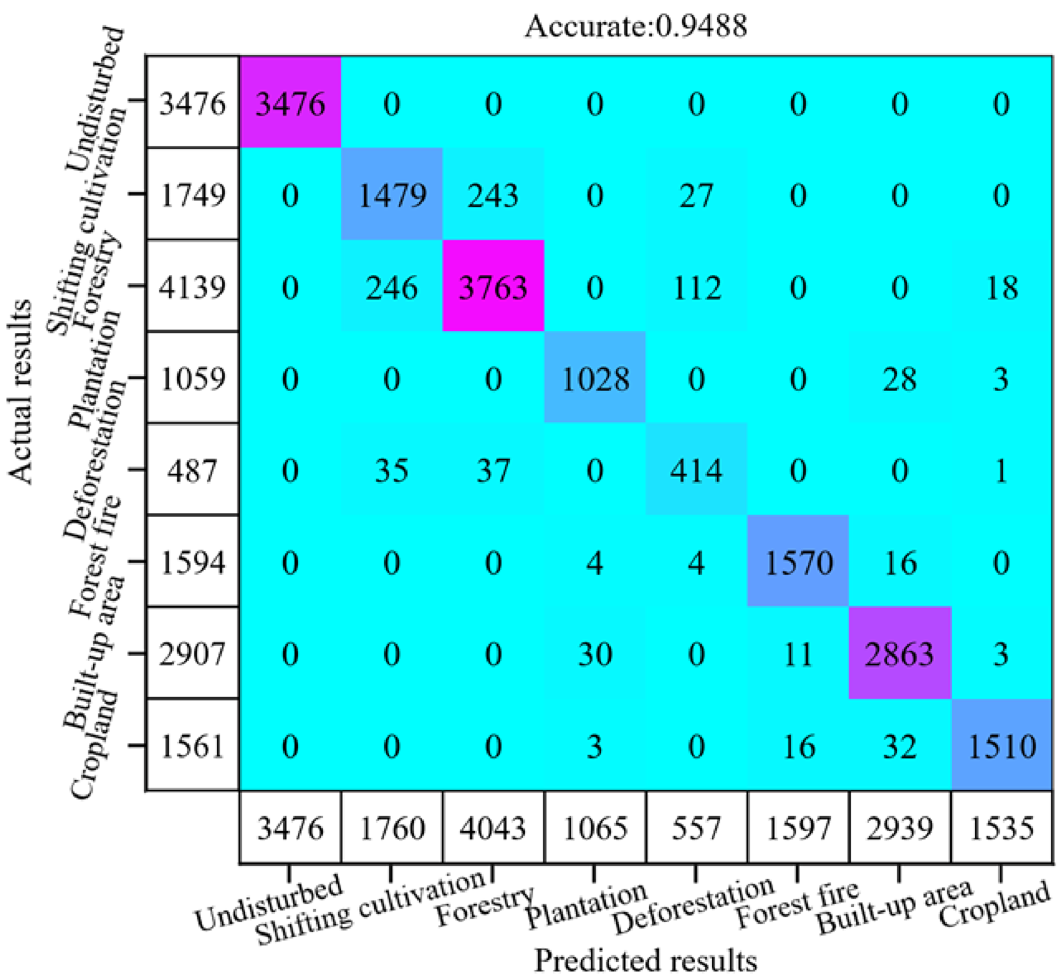
$$V(P_j) = \frac{P_j(1 - P_j)}{n_{.j} - 1} \tag{7}$$

Comment #4

c) Row and column totals need to be added to Figure 3. (**Revised**)

Response #4

Thanks for your suggestion. We have added the total of rows and columns in Figure 3 as per your suggestion (Page 12, Figure 3):



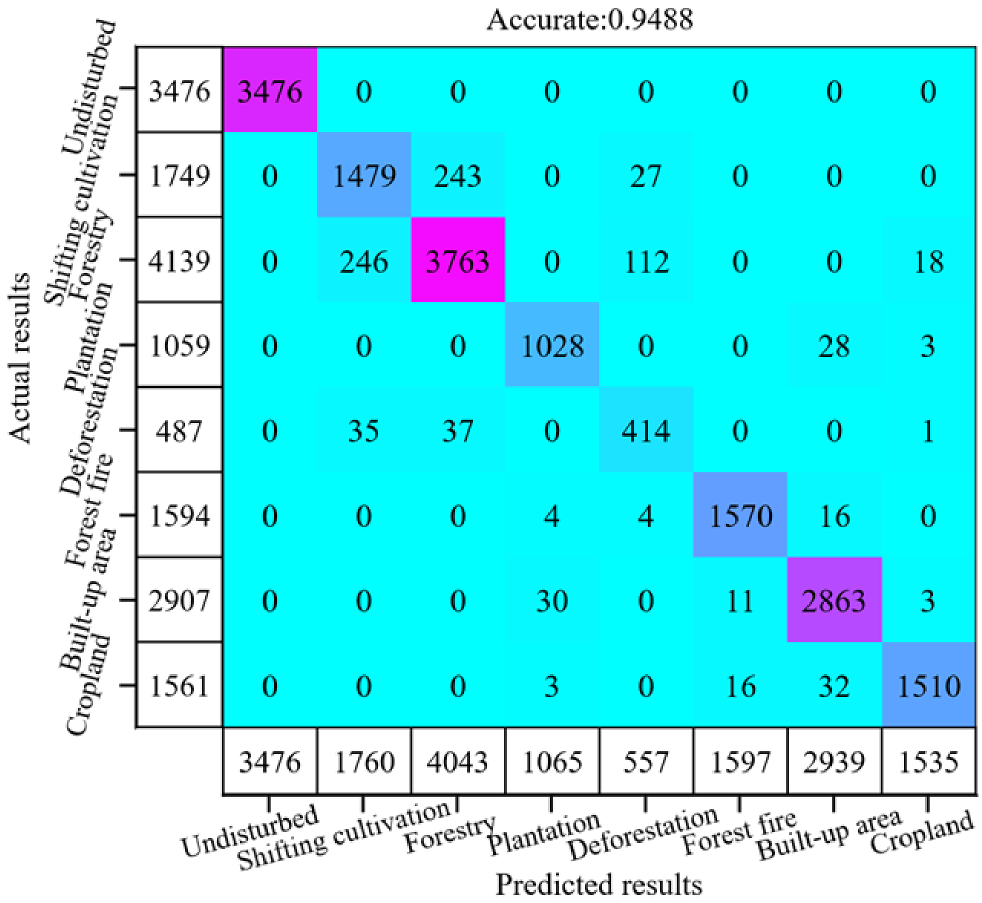
"Figure 3 Confusion Matrix of Global Forest Disturbance Classification"

Comment #5

d) It is unclear what the vertical color bar on the right of the figure represents (range from 0 to 40,000). Please remove it or explain what it is. (**Revised**)

Response #5

Thanks for your suggestion. We have removed the vertical color bar in Figure 3 as per your suggestion (Page 12, Figure 3):



"Figure 3 Confusion Matrix of Global Forest Disturbance Classification"

Comment #6

e) I will identify this comment as purely an opinion, but I am skeptical that a disturbance product can achieve the high accuracies reported. Accurately mapping forest change is exceedingly difficult, so to achieve user's and producer's accuracies of over 95% for many of these disturbance types doesn't seem possible. Comment 2a is related to this same concern. **(Revised)**

Response #6

Thanks for your comment. We acknowledge that accurately mapping forest disturbance types remains highly challenging. We also note that for certain complex disturbance types, such as shifting cultivation and deforestation, our accuracy is even less than 80%. Higher accuracy was achieved for types such as plantations and fires, which benefits from significant advances in global single-type disturbance mapping efforts, including existing plantation and forest fire map, etc. In our identification process, these relatively accurate auxiliary datasets were incorporated as features in the machine learning model, thereby improving accuracy for these *categories. Furthermore, permanently changed areas—such as forests converted permanently to cropland or built-up areas—also showed higher accuracy, owing to the use of highly reliable land cover datasets from around 2020. These auxiliary datasets substantially supported the identification of certain disturbance types. We have now enhanced the manuscript with detailed descriptions of the sources and applications of these datasets.*

In Study workflow section (Page 4, Line 89-92 and Table2):

"These features were systematically derived from both temporal and spatial dimensions, including: …… disturbance potential metrics (DP), land use/cover features (LUC)……"

Table 2 Global Forest Disturbance Characteristics Indicator

Indicator type	Forest disturbance characteristic indicators		
…	…	…	…
DP	Forest fire area	Plantation area	Intensity of population
LUC	2020 Land Use /Cover	Forest cover in 2000	Forest cover in 2020
…	…	…	…

In Land use/cover dataset section (Page 6, Line 120-129 and Table3):

"To assess the mapping accuracy of global forest change areas, this study incorporated multiple authoritative land cover and forest cover products as reference datasets, including: (1) ESA WorldCover 2020,….. identifying 2020 forest cover distribution, and classifying non-forest land cover types in 2020. The detailed information of these datasets is systematically documented in Table 3."

Table 3: Source of Land Cover Dataset

Dataset	Resolution	Dataset source and main purpose
ESA WorldCover 2020	10m	Used to assist in identifying disturbances in cropland, built-up areas, etc. https://worldcover2020.esa.int/
…	…	…

In Ancillary datasets section (Page 8, Line 157-164):

"Forest disturbance has strong disturbance sources. Therefore, using existing disturbance source datasets to assist in identifying typical forest disturbance types can effectively improve the accuracy of mapping results. Generally, a high spatial consistency is typically observed between disturbance types such as forest fires and plantation expansion and global fire and plantation distribution. This study uses global fire distribution datasets, artificial plantation distribution datasets, oil palm datasets, and other auxiliary methods to identify forest disturbance types. Meanwhile, there is a high correlation between population distribution and

forest disturbance. This study collected a forest disturbance potential dataset from three aspects: population density, forest fire distribution, and spatial distribution of oil palms."

In Result section (Page 12, Line 267-268):

"Owing to the high accuracy of current DP and LUC datasets, such as forest fire maps, plantation maps, and land use/cover maps, the identification accuracy of their corresponding disturbance types is considerably improved."

Comment #7

The accuracy estimates reported on page 12 and in Table 4 are also a cause for concern.

a) It is evident that the stratified formulas were not used to estimate producer's accuracy and overall accuracy. If the sampling design is stratified and the stratified formulas were not used, these estimates would be incorrect. (**Revised**)

Response #7

Thanks for your comment. Thank you very much for your comment. We have revised the formula expression in the paper and re-evaluated the accuracy estimates (Page 13, Line 276-286 and Table4):

"The accuracy assessment results reveal significant variations in classification performance across different forest disturbance types. The overall accuracy reaches 94.88% (±0.17%), indicating robust model performance at the aggregate level (Table 4). Forest fire disturbance (98.31% ±0.32% user's accuracy, 98.49% ±0.31% producer's accuracy) and cropland occupation (98.37% ±0.32%, 96.73% ±0.45%) demonstrate the highest classification reliability. Forestry replanting shows strong(93.07% ±0.40%, 90.92% ±0.45%), while shifting cultivation achieves moderate performance and slightly more variable accuracy (84.03% ±0.87%, 84.56% ±0.86%). Deforestation of natural forests exhibits the lowest user's accuracy (74.33% ±1.85%), suggesting significant confusion with other disturbance types, despite its relatively higher producer's accuracy (85.01% ±1.62%). Built-up area expansion shows nominally high accuracy (97.41% ±0.29%). These results highlight the model's effectiveness for dominant disturbance types."

Table 4 Accuracy Evaluation of GFD Mapping Results

Type	User 's Accuracy	Uncertainty (±)	Producer's Accuracy	Uncertainty (±)	Overall Accuracy
11	84.03%	0.87%	84.56%	0.86%	94.88%± 0.17%
12	93.07%	0.40%	90.92%	0.45%	

13	96.53%	0.56%	97.07%	0.52%
14	74.33%	1.85%	85.01%	1.62%
15	98.31%	0.32%	98.49%	0.31%
18	97.41%	0.29%	98.49%	0.23%
19	98.37%	0.32%	96.73%	0.45%

Comment #8

b) It seems very likely that the standard error values are incorrect for several cases. For example, if we had a simple random sample with a sample size of n=17,000 (approximate sample size of matrix in Figure 3), the standard error of overall accuracy would be SQRT[(0.95)*(0.05)/17000]=0.0033 or 0.33%. The reported standard error for overall accuracy is 2.86% from line 253, nearly 10 times larger. The standard errors for producer's accuracy of Types 18 and 19 (approximately 20% and 15%) are suspiciously large given the large sample sizes for these two disturbance types. Lastly, the standard errors reported for user's accuracy also don't match what I calculate if I apply equation (7) to the data in Figure 3. Please re-check the standard error estimates to confirm. (**Revised**)

Response #8

Thanks for your comment. Thank you very much for your comment. We have revised the formula expression in the paper and re-evaluated the accuracy estimates (Page 13, Line 276-286 and Table4). The latest results are as you calculated, and there is a significant improvement in the variance of various accuracy metrics. The standard error of overall accuracy is SQRT[（0.9488）（1-0.9488）/16972]= 0.17%.*

"The accuracy assessment results reveal significant variations in classification performance across different forest disturbance types. The overall accuracy reaches 94.88% (±0.17%), indicating robust model performance at the aggregate level (Table 4). Forest fire disturbance (98.31% ±0.32% user's accuracy, 98.49% ±0.31% producer's accuracy) and cropland occupation (98.37% ±0.32%, 96.73% ±0.45%) demonstrate the highest classification reliability. Forestry replanting shows strong(93.07% ±0.40%, 90.92% ±0.45%), while shifting cultivation achieves moderate performance and slightly more variable accuracy (84.03% ±0.87%, 84.56% ±0.86%). Deforestation of natural forests exhibits the lowest user's accuracy (74.33% ±1.85%), suggesting significant confusion with other disturbance types, despite its relatively higher producer's accuracy (85.01% ±1.62%). Built-up area expansion shows nominally high accuracy (97.41% ±0.29%). These results highlight the model's effectiveness

for dominant disturbance types."

Table 4 Accuracy Evaluation of GFD Mapping Results

Type	User 's Accuracy	Uncertainty (±)	Producer's Accuracy	Uncertainty (±)	Overall Accuracy
11	84.03%	0.87%	84.56%	0.86%	94.88%± 0.17%
12	93.07%	0.40%	90.92%	0.45%	
13	96.53%	0.56%	97.07%	0.52%	
14	74.33%	1.85%	85.01%	1.62%	
15	98.31%	0.32%	98.49%	0.31%	
18	97.41%	0.29%	98.49%	0.23%	
19	98.37%	0.32%	96.73%	0.45%	

Comment #9

c) Note that Type 18 in Table 4 is accidentally mis-labeled as "118" (**Revised**)

Response #9

Thanks for your suggestion. We have revised the expression here (Page 13, Table 4):

Table 4 Accuracy Evaluation of GFD Mapping Results

Type	User 's Accuracy	Uncertainty (±)	Producer's Accuracy	Uncertainty (±)	Overall Accuracy
11	84.03%	0.87%	84.56%	0.86%	94.88%± 0.17%
12	93.07%	0.40%	90.92%	0.45%	
13	96.53%	0.56%	97.07%	0.52%	
14	74.33%	1.85%	85.01%	1.62%	
15	98.31%	0.32%	98.49%	0.31%	
18	97.41%	0.29%	98.49%	0.23%	
19	98.37%	0.32%	96.73%	0.45%	

Comment #10

Table 5 provides estimates of area of the GFD types. Presumably these are from the inadequately described "validation" sample. The Abstract should be revised to clarify what is presented in the manuscript. The manuscript's title suggests that the primary purpose of the manuscript is to present a new global forest disturbance dataset (i.e., a map). But key parts of the manuscript are sample-based estimates of area, which would use the disturbance map for stratification, but the key data are then the sample and disturbance type labels provided by the expert interpreters. For area estimation the role of the new disturbance map is secondary. If the main objective of the manuscript is to provide this global dataset, then sample-based area estimates would seem unnecessary and only the accuracy results would be necessary to present.

This same ambiguity is present in the Conclusion section. Lines 354-358 highlight the map of disturbance. But without any transition flagging the use of sample-based area estimation, Lines 358-360 then report sample-based estimates of area (Table 5) that use only the map through stratification of the sample. Please revise the Abstract and Conclusion to more clearly identify the purpose of the map and the role of sample-based area estimation to the objectives of the manuscript. (**Revised**)

Response #10

Thank for your suggestion. We have removed unnecessary area estimation content and Table 5 as per your suggestion to highlight the focus of our research - the GFD map dataset. The abstract and conclusion also removed the ±% area estimation section (Page 1, Line 15-19; Page 19, Line 362-364):

"The results reveal that forestry replanting (43.79), shifting cultivation (24.32%), and forest fires (11.45) dominate global forest loss. There are regional differences in global forest disturbance, such as farmland expansion in South America and Africa, forest fires in northern regions, and shifting cultivation in tropical regions. Disturbed forests span 1,247.06Mha, accounting for 30.87% of the global forest area. Notably, 2.76% of global forests were newly established, primarily in China, India, and Brazil."

"The results highlight forestry replanting (43.79%), shifting cultivation (24.32%), and forest fires (11.45%) as the dominant drivers of global forest cover changes, collectively accounting for nearly 80% of the total disturbed area. Meanwhile, the newly added forests worldwide account for 2.76% of the global forest disturbance area."

Technical Corrections

Comment #11

1. Line 15: It is not clear whether the number to the right of the +/- is a standard error or a margin of error of a confidence interval. Please identify more clearly. (**Revised**)

Response #11

Thanks for your comments. This is the confidence interval. Based on your previous suggestion (Comment #10), we have removed the ± percentage related to area estimation.

Comment #12

2. Lines 19-20: The comparison to other datasets provides an evaluation of "agreement" or "consistency" with these other datasets. These other datasets are not "truth". Therefore,

agreement with these other datasets does not "confirm reliability" or convey "accuracy" but instead quantifies consistency with other datasets. (**Revised**)

Response #12

Thanks for your suggestion. We have revised the expression here (Page 1, Line 19-20):

"The spatial consistency analysis (R^2 = 0.93) highlights a strong overall agreement between the GFD product and other datasets, while the GFD product offers superior spatial resolution."

Comment #13

3. Line 43: What specifically is "subjective" about field surveys? The implication is that remote sensing is not subjective, but that would seem dubious because surely there are subjective components of remote sensing as well. (**Revised**)

Response #13

Thanks for your suggestion. We have revised the expression here (Page 2, Line 44-46):

"Traditional forest monitoring predominantly relies on field surveys, which are limited by low temporal resolution and high labor costs, rendering them difficult to scale for large-area or frequent-assessment applications (Scheeres et al., 2023; Finger et al., 2021)."

Comment #14

4. Lines 19, 225, 226, 321: This is a minor point, but stating that a comparison is made with "existing" datasets is not meaningful because we obviously cannot make a comparison to a dataset that does not exist. It would be better to use "other datasets" instead of "existing datasets". (**Revised**)

Response #14

Thanks for your suggestion. We have addressed all similar issues in the manuscript (such as Page 11, Line 245-247):

"2.5.3 Comparison with other datasets

We compared GFD map with these other datasets to calculate the degree of agreement on typical forest disturbance types."

Comment #15

5. Page 10, equation (10): This formula for the standard error of the estimated proportion of area does not match equation (10) presented in Olofsson et al. (2014). (**Revised**)

Response #15

Thanks for your comment. Based on your previous suggestion (Comment #10), we have removed related content.

Comment #16

6. Equation (11): The use of "UA" for the standard error will be confusing because it could easily be misread as an abbreviation for "User's Accuracy" and "UA" provides no obvious connection to standard error. (**Revised**)

Response #16

Thanks for your comment. Based on your previous suggestion (Comment #10), we have removed related content.

Comment #17

7. Equation (12): Please check this formula. It seems unlikely that there would be a "bar" above qi (indicating a mean) in the denominator but no "bar" above pi in that same denominator. (**Revised**)

Response #17

Thanks for your suggestion. We have revised this formula according to your suggestion (Page 11, Line 260).

$$R^2 = 1 - \frac{\sum_i^n (q_i - p_i)^2}{\sum_i^n (q_i - \overline{p_i})^2} \tag{8}$$

Comment #18

8. Line 226: Because these other datasets are not "truth", comparisons to these datasets would represent "agreement" and "disagreement". Use of the term "errors" does not seem appropriate here. (**Revised**)

Response #18

Thanks for your suggestion. We have revised the expression here (Page 11, Line 246-247):

"We compared GFD map with these other datasets to calculate the degree of agreement on typical forest disturbance types."

Comment #19

9. Line 234: a space should be inserted between "s" and "p" in "asp". (**Revised**)

Response #19

Thanks for your suggestion. We have revised the expression here (Page 11, Line 254-255):

"……each grid relative to the global forest loss area, denoted as p_i, where i represents the driver type……".

Comment #20

10. Table 4: state what the +/- columns represent. (**Revised**)

Response #20

Thanks for your suggestion. We have added corresponding explanations in Table 4 (Page13, Table4). The ± represent uncertainty of accuracy.

Table 4 Accuracy Evaluation of GFD Mapping Results

Type	User 's Accuracy	Uncertainty (±)	Producer's Accuracy	Uncertainty (±)	Overall Accuracy
11	84.03%	0.87%	84.56%	0.86%	94.88%± 0.17%
12	93.07%	0.40%	90.92%	0.45%	
13	96.53%	0.56%	97.07%	0.52%	
14	74.33%	1.85%	85.01%	1.62%	
15	98.31%	0.32%	98.49%	0.31%	
18	97.41%	0.29%	98.49%	0.23%	
19	98.37%	0.32%	96.73%	0.45%	

Comment #21

11. Line 288: The meaning of "robust" precision is unclear. In what sense can precision be "robust"? (**Revised**)

Response #21

Thanks for your comment. Based on your previous suggestion (Comment #10), we have removed this content related to area estimation.

Comment #22

12. Line 326: "MEA" should be "MAE" and the word "only" should be removed from before "13%" as that is a value judgment of magnitude of the disagreement. (**Revised**)

Response #22

Thanks for your suggestion. We have revised the expression here (Page 16, Line 334-335).

"From the perspective of error, the MAE and RMSE of the two are 13% and 19%, respectively (Fig. 8)."

Comment #23

13. Panel b) of Figure 8 should be deleted or perhaps converted to a small table. The R^2, MAE, and RMSE values do not make much sense for only 6 data points and the "All Types" case must have a massive influence on the summary statistics. (**Revised**)

Response #23

Thanks for your suggestion. We have removed Figure 8b as per your suggestion (Page 17, Figure 8).

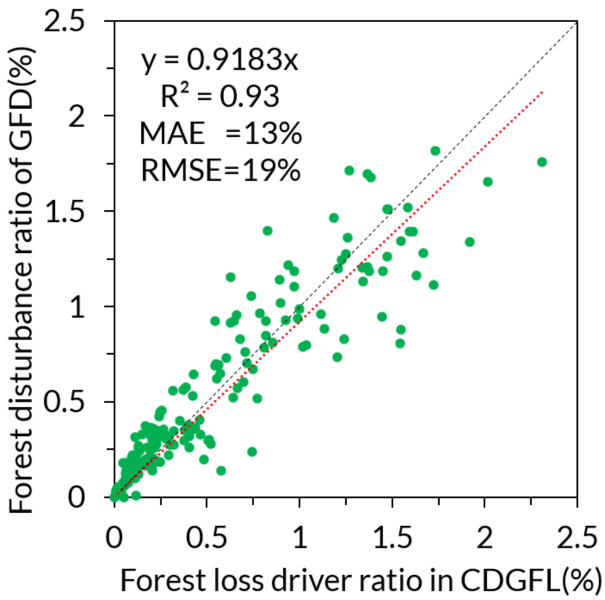


Figure 8 Overall spatial consistency comparison with CDGFL.

Comment #24

14. Lines 338-340: "MEA" should be "MAE" in multiple places. (**Revised**)

Response #24

Thanks for your suggestion. We have revised the expression here (Page 17, Line 342-346).

"The comparative analysis reveals that these four major disturbance types display high spatial agreement with the existing low-resolution CDGFL dataset, with the following metrics: shifting cultivation (R^2=0.78, MAE=6.76%, RMSE=15.71%), forestry replanting (R^2=0.83, MAE=10.61%, RMSE=17.49%), forest fire (R^2=0.85, MAE=5.93%, RMSE=12.17%), and deforestation of natural forest (R^2=0.62, MAE=4.66%, RMSE=11.47%)."

Comment #25

15. Throughout the manuscript the word "samples" is used incorrectly. The definition of "sample" in statistics is that it is a subset of n units selected from the population. The individual elements of that sample are "sample units", in this case a sample unit is a 30-m pixel. Thus, there are not 57,000 "samples" (e.g., Line 13), but one "sample" consisting of 57,000 sample units or sample pixels. This incorrect use of "samples" should be corrected throughout the manuscript. (**Revised**)

Response #25

Thanks for your suggestion. We fully agree with your opinion. Due to the fact that our sample contains 57000 sample points. Therefore, we have replaced "samples" with "sample points" throughout the manuscript (such as Page 14, Line 86; Page 41, Line 176-179).

" The model training and validation incorporated 57,000 expertly labeled sample points of forest disturbance……"

"Although these pre-labeled sample points are uncertainty, they provide our interpreters with rapid and spatially randomized regions for sample point selection. Subsequently, our interpreters performed visual verification and interpretation to mark accurate sample points within the vicinity of these pre-labeled locations."