Editor

Dear Franziska Müller,

We are pleased to inform you that the editor report for the following NHESS manuscript is now available:

egusphere-2025-4880
Title: Hybrid forest disturbance classification using Sentinel-1 and inventory data: a case-study for Southeastern USA
Author(s): Franziska Müller et al.
MS type: Research article
Iteration: Major revision

The editor has decided that Major revisions are necessary before the review process can be continued. Please log in using your Copernicus Office user ID 684410 to find the editor report at:
https://editor.copernicus.org/NHESS/ms_records/egusphere-2025-4880

We kindly ask you to revise your manuscript accordingly and to upload the revised files, a point-by-point reply to the comments, and a marked-up manuscript version showing the changes made no later than 25 Apr 2026 at:
https://editor.copernicus.org/NHESS/review-file-upload/egusphere-2025-4880

Please find all information on manuscript submission at:
https://www.natural-hazards-and-earth-system-sciences.net/for_authors/submit_your_manuscript.html

Your revised manuscript will be reviewed again and you will be informed about the outcome by separate email.

Besides adjustments requested by the editor or referees, please check your manuscript carefully for typos, missing co-authors and their affiliations, terminology, updates of data in tables, or updates of variables in equations. All these have to be clarified with the editor and therefore have to be included before you submit your revised manuscript. Should your manuscript be finally accepted it will not be possible to include such rather substantial changes anymore when your manuscript is in final production (proofreading).

Please note that all referee and editor reports, the author's response, as well as the different manuscript versions of the peer-review completion (post-discussion review of revised submission) will be published if your paper is accepted for final publication in NHESS.

You are invited to monitor the processing of your manuscript via your MS overview at:
https://editor.copernicus.org/NHESS/my_manuscript_overview

In case any questions arise, please do not hesitate to contact me. Thank you very much for your cooperation.

Kind regards,

The editorial support team

Dear Editor,

Thank you for handling our manuscript and for the constructive feedback from the reviewers. We sincerely appreciate the opportunity to improve our work.

We have carefully revised the manuscript in response to the comments and suggestions provided. In particular we have added two new sections to the Discussion: "Methodological Assumptions and Limitations" and "Generalizability and Applications of the Proposed Approach", which provide a more comprehensive evaluation of the methodological framework and its broader relevance. We have clarified the change detection approach, including a brief threshold analysis to support the selection of key parameters.
In the Results section, we have added one new figure and one new table to better illustrate the performance of the hybrid classification framework and its sensitivity to input parameters. We have carefully proofread the revised manuscript to correct any remaining grammatical, typographical, or stylistic issues.

Thank you again for your time and consideration.

Sincerely,

Franziska Müller on behalf of the co-authors

Hybrid forest disturbance classification using Sentinel-1 and inventory data: a case-study for Southeastern USA

Franziska Müller, Laura Eifler, Felix Cremer, Pieter Beck, Gustau Camps-Valls, and Ana Bastos, *EGUsphere*

Response to Reviewer #1

The paper combines three different datasets on forest disturbances for a large region in the Southern and Eastern US: IDS, S1DM, and Planetscope. The first one is a routine product freely accessible with yearly updates, the second one is generated by the authors based on Sentinel-1 time series, and the third one is merely a case-based evaluation tool including manual delineation of polygons with damaged forest.

R1C1: Out of the many disturbances, the authors select windthrow, bark beetle attacks, and defoliators, representing important categories, but leaving out fire.

We thank the reviewer for highlighting this point. We agree that fire is indeed a critically important disturbance category. For example, Global Forest Watch reports that between 2001 and 2024, approximately 151 Mha of tree cover loss worldwide was attributed to fire, roughly the size of Mongolia.

At the same time, we chose not to include fire in our study because there are already excellent, globally available fire-monitoring systems, such as NASA's FIRMS (Fire Information for Resource Management), which provide near-real-time detection of active fires from MODIS and VIIRS, often within 3 hours of satellite observation. These resources offer a level of observation and detection that is already outstanding and widely used.

Our study, therefore, focuses on disturbances that are less well monitored by remote sensing, such as windthrow, bark beetle outbreaks, and defoliators, for which detection challenges are known (Mcdowel et al., 2015; Viana-Soto & Senf, 2025; Schroeder et al., 2017). Therefore, we believe our approach can provide novel insights and added value.

R1C2: Which areas are considered and which are excluded is largely determined by the flight trajectories leading to the IDS dataset - this is a bit disappointing since S1 tiles have global coverage, and the advantage of satellite remote sensing is the ability to conclude on patterns outside ground-based (or aerial for that matter) observations, provided sufficient training data. It would be interesting to see an area with a damage according to the Planetscope manual delineation and the S1 trend detection there, but not covered by the IDS. Of course, the TCC might still be used to exclude non-forested areas. The exercise would demonstrate the power of the S1 disturbance detection completely independent from the IDS data; in general, the two approaches are just compared, they do not depend on each other.

We thank the reviewer for this insightful comment. Indeed, Sentinel-1 has global coverage, and in principle, S1-based disturbance detection could be applied independently of the IDS dataset; combined with TCC, this could provide valuable insights into forest disturbances.

However, the reviewer's suggestion would require identifying disturbed patches based on Planet data that are not detected by IDS. This would imply that other high-quality, spatially explicit, disturbance classification datasets independent of IDS would exist. This is precisely the knowledge gap we hope to address in the future through satellite-based disturbance detection and classification.

Detailed information on disturbance agents, especially insects, is available in the IDS dataset and, to some extent, in the Forest Inventory Assessment, but it is based on plot-level sampling with irregular revisit times (see Eifler et al., 2026).

This is why our analysis is limited to IDS-covered areas: without IDS, we would be unable to associate observed disturbances with their causal agents.

We agree that such an exercise could be valuable and highlight it as an interesting avenue for future research in our manuscript.

> L.682-690 (in 5.3. *Methodological Assumptions and Limitations*): *In this study, our analysis was constrained to areas covered by the IDS dataset, which provides detailed, spatially explicit information on disturbance agents. While Sentinel-1 offers global coverage and the potential to detect disturbances beyond regions surveyed by aerial or ground-based inventories, the lack of independent reference data limits our ability to confidently attribute detected disturbances to specific causal agents in areas outside IDS coverage. Consequently, the current framework cannot fully evaluate the performance of Sentinel-1-based disturbance detection across unsurveyed regions or in landscapes where IDS data are absent. Future work should address these limitations by integrating Sentinel-1 disturbance detection with high-resolution, independent reference datasets, such as manual delineations from Planet imagery, and by using tree canopy cover products to exclude non-forested areas. Expanding the analysis in this way would allow for a more comprehensive assessment of Sentinel-1's capabilities and its potential for large-scale, autonomous forest disturbance monitoring.*

R1C3: Concerning the exclusion of non-forest areas, the authors set an unnecessary strict threshold for the presence of a forest, i.e. 30% canopy cover. This is not aligned with the FAO definition of a forest as any are of minimum size 0.5 ha with a canopy cover (for trees which can grow to more than 5 meters) of only 10% (https://fra-data.fao.org/definitions/fra/2020/en/tad). As the minimum area required according to FAO is only 12.5 pixels for S-1, the 30% seems to be overly restrictive.

We thank the reviewer for this thoughtful comment.

Given the FAO definition, a minimum canopy cover of 10% for a minimum size of 0.5ha corresponds to ca. 500m^2, which is larger than the 400m^2 (20x20m) of the S-1 pixels considered here. The main reason to exclude pixels with low Tree Canopy Cover (TCC) is, however, the nature of the S-1 signal. At the scale of an S-1 pixel, a large fraction of non-forest vegetation types (e.g., croplands with strong seasonal phenology) could influence the change-detection algorithm, for reasons not necessarily associated with forest disturbances. Therefore, our strict filtering aims to ensure that the patches identified as “disturbed” are associated with forest disturbances rather than other types of interventions, such as agricultural practices. Nevertheless, we reran the whole analysis using a less restrictive TCC threshold of 10% as suggested. The results remain qualitatively very similar to those presented in the manuscript, as shown in the following result figures.

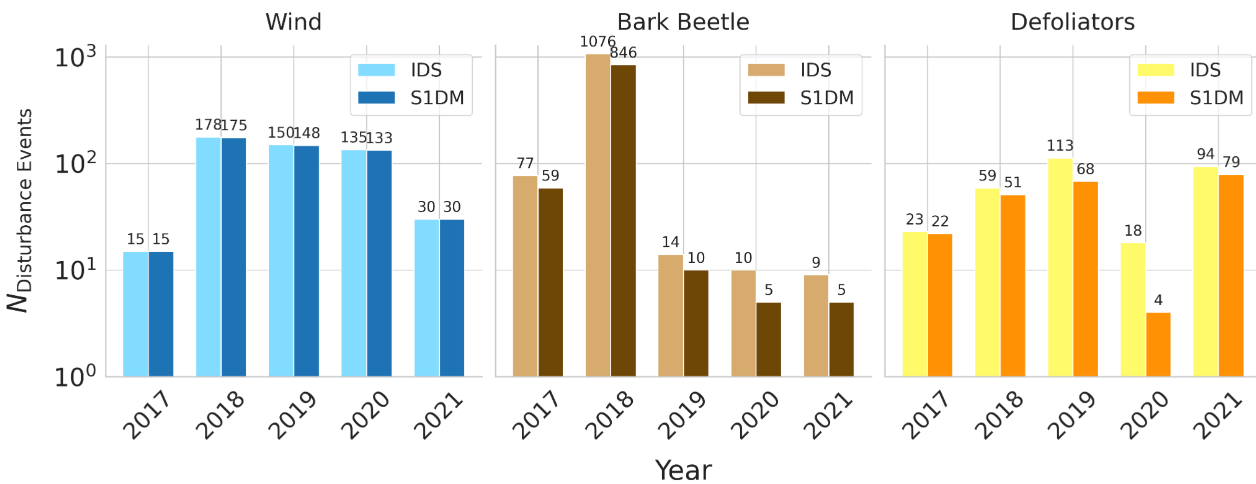

Figure R.1.1: Detection efficiency of various disturbance types using Radar Change Detection with the TCC threshold of 10%. The bar plot shows the number of events on the y-axis (log scale) for each disturbance type on the x-axis, with IDS shown in light color and S1DM in darker color. The total number of IDS disturbance events that had a corresponding S1DM signal within the 500 m buffer was 501 for wind (S1DM: 501; IDS: 508), 925 for bark beetle (S1DM: 925; IDS: 1186), and 224 for defoliators (S1DM: 224; IDS: 307).

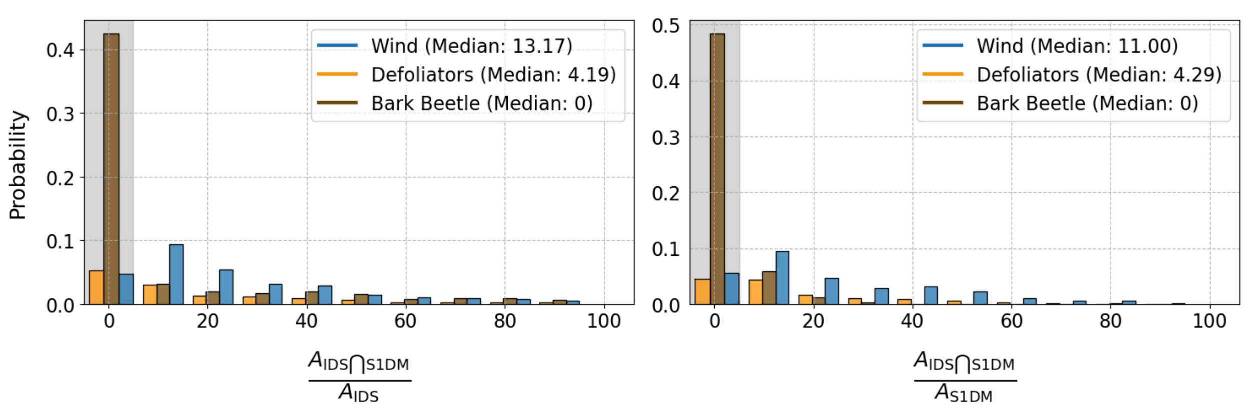

Figure R1.2: Probability of (a) IDS polygons being covered by S1DM, and (b) S1DM polygons overlapping with IDS, for each disturbance type. Disturbance types are color-coded as follows: Wind (blue), Bark Beetle (brown), and Defoliators (yellow). Median probabilities for each disturbance type are indicated in parentheses in the legend.

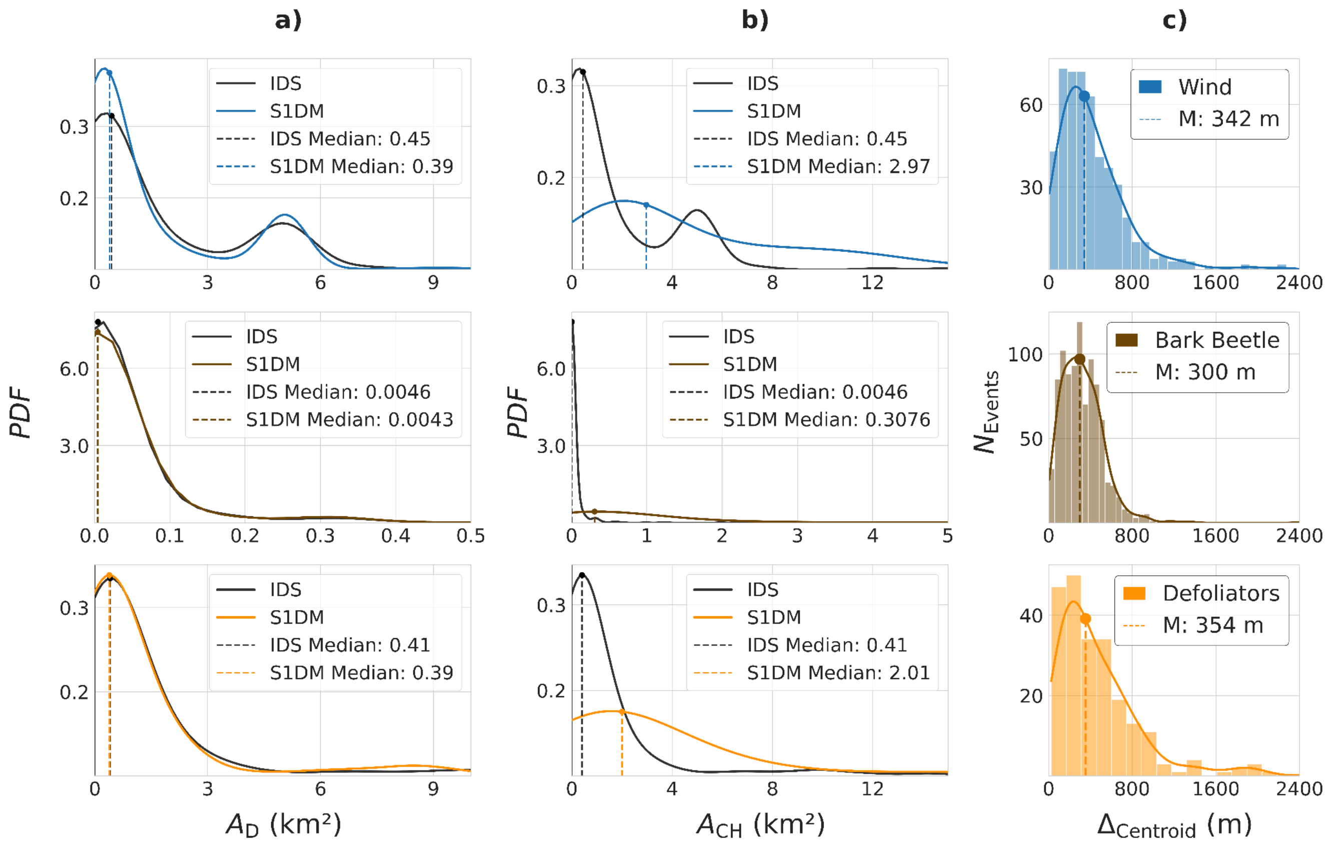

Figure R1.3: Comparison of disturbance size (in km2) and spatial bias (in m) between the IDS and S1DM datasets tested on forested areas with a TCC threshold of 10%. The figure consists of three vertical panels, each displaying the results for three disturbance types: Wind (top, blue), Bark Beetle (second, brown) and Defoliators (third, green). a) shows the probability density function (PDF) of disturbance patch areas for IDS (black lines) and S1DM (colored lines) across the four disturbance types. b) presents the PDF of the convex-hull (CH) areas, representing the spatial spread of the disturbance polygons, using the same color scheme for S1DM and IDS as in the left panel. c) displays histograms and corresponding PDFs of the spatial distance between IDS and S1DM centroids for each disturbance type.

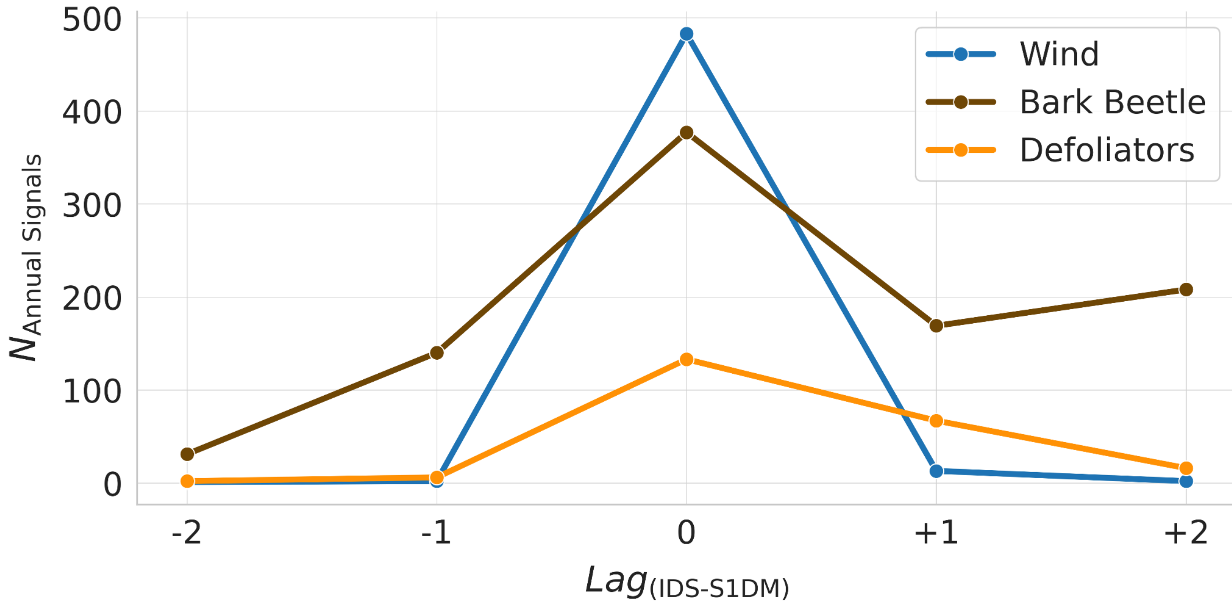

Figure R1.4: Temporal relationship between disturbance detection by S1CD and IDS with a TCC threshold of 10%. The x-axis represents the temporal yearly agreement with IDS - S1CD. Negative values indicate that IDS recorded the disturbance before S1CD detected it. The y-axis shows the count of events corresponding to each year's lag, with different disturbance types represented by distinct colored lines. We focus on the single S1DM timestamp closest to the IDS year for each event. We calculate the temporal lag by subtracting the S1DM detection year from the IDS year. Negative values indicate that S1DM detection occurred after IDS, while positive values indicate earlier radar detection.

With the 10% TCC threshold, the total number of detected IDS events increases slightly (Bark Beetle: from 1,177 to 1,186; Wind: from 478 to 508; Defoliators: from 213 to 307). A similar increase is observed for the corresponding S1CD events (Bark Beetle: from 899 to 925; Wind: from 469 to 501; Defoliators: from 144 to 224). Importantly, despite these changes in absolute numbers, the relative loss percentages remain largely consistent: Bark Beetle changes from −23.62% to −22.01%, Wind from −1.88% to −1.38%, and Defoliators from −32.39% to −27.04%.

Overall, this analysis confirms that our results are robust to a less restrictive TCC threshold. However, for the main analysis, we retain the 30% TCC threshold, as it provides a more conservative and reliable estimate of forested areas.

We have added a clarification to the manuscript in Section 3.2 (Tree Canopy Cover):

> L. 290-295: "We applied a 30% TCC threshold, which is stricter than the FAO forest definition (≥10% canopy cover; FAO, 2020), to ensure that detected disturbed patches correspond to forest disturbances and not changes in vegetation structure driven by other processes, such as agricultural practices. Using a 10% threshold yielded very similar results, with fewer than 80 additional events per disturbance type."

R1C4: On page 3, l. 84-92, ML and DL are mentioned. While it is largely correct what is written here, it seems that not a single ML or DL method is applied in the rest of the paper. What is the purpose of that paragraph? Is it a leftover from earlier versions of the manuscript? It might as well be deleted.

We agree with the reviewer that no machine learning or deep learning model is applied in this manuscript. However, the point we make in the introduction (which is then discussed later in the manuscript) is that these methods have been increasingly used to advance forest disturbance mapping, but they are data-intensive and require reliable labels, which in turn are labor-intensive and costly. Here we show that while IDS provides unique information about disturbance agents, it also suffers from limitations regarding the spatial delineation of the disturbed patches, as exemplified in Figures R1.5, R1.7. Also, see a more in-depth analysis of the spatio-temporal uncertainty of IDS data in Eifler et al. (2026). Our goal is thus two-fold: improve the delineation of disturbed patches identified by IDS using radar information from S-1 and, while doing so, produce a standardized benchmark for training and validating ML and DL models.

Since the detailed discussion of DL/ML takes up considerable space in the introduction but is not central to the manuscript's topic, we have shortened the paragraph as follows:

> L.88-93: "Increasingly, remote sensing studies use data-driven approaches (e.g., Artificial Intelligence) to detect, map, and classify forest disturbances, demonstrating their ability to capture complex spatiotemporal patterns (Andresini et al., 2024; Bárta et al., 2021; Hawryło et al., 2018; Gibson et al., 2020). The effectiveness of these models depends on high-quality reference data, as accurate ground-truth labels are essential for training and validating disturbance detection algorithms. This data-hungry approach creates a bottleneck because labeling forest disturbances requires extensive labor and expertise to distinguish agents such as insects, drought, or biotic stress."

R1C5: Regarding the size of disturbed areas, a maximum size of 15 km2 is used. This seems to be an arbitrary threshold and a huge difference to the maximum size of 2000 km2 used by Eifler et al. (2024). The only justification is (l. 278) "we applied a stricter filter". Certainly you did, but why? Many beetle attacks happen on or spread to larger areas, similar with defoliators.

We thank the reviewer for this important comment. We agree that the 15 km² maximum disturbance threshold may currently appear arbitrary, especially compared to the 2000 km² threshold used by Eifler et al. (2026).

To provide context, we analyzed the distribution of disturbance sizes in our dataset, focusing on bark beetles, defoliators, and wind. Out of 27,776 disturbances (before excluding compound events), only 257 events (~0.93%, i.e., the top 1%) exceeded 15 km², while the vast majority, 27,519 events (~99.07%), were smaller or equal to this threshold. The mean area of the larger disturbances was 120.67 km².

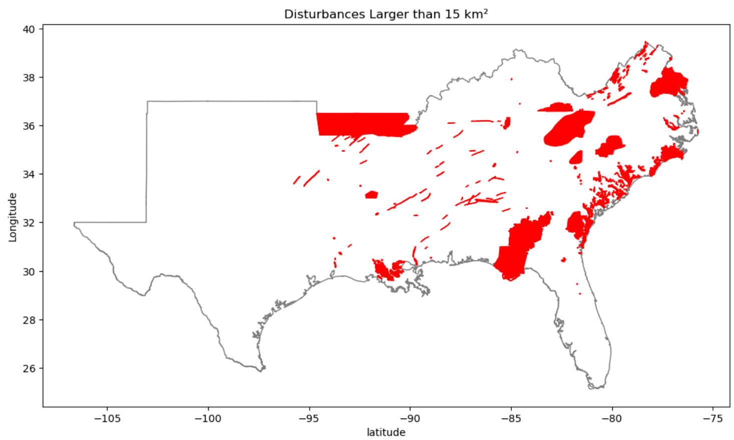

Figure R1.5: Disturbances (bark beetles, defoliators, and wind) before preprocessing with sizes larger than 15 km² in the study region.

This size, combined with visual inspection (in the provided Figure R1.5 above), indicates that these unusually large disturbances are likely the result of human error/imprecision rather than true ecological events. Given that these are outliers in terms of extent and likely unreliability, we applied the stricter 15 km² filter to focus our analysis on realistically sized disturbance events while avoiding potential bias.

To clarify this point, we have changed the following sentence to explain our decision of excluding the above 15 km² elements:

> L. 271-275: "**However, such cases were extremely rare, with fewer than 1%** of polygons exceeding **15 km². Despite their rarity, these large polygons represented very extensive and poorly defined disturbance areas (mean size ~121 km²), and we removed them to avoid introducing bias and uncertainty into our analysis. While Eifler et al. (2026) excluded only** events larger than 2000 km², we applied a stricter filter **by excluding** polygons larger than 15 km²."

R1C6: The polygon sizes mentioned in Table 1 are hard to believe. The smallest one (0.5 m2) would cover a single tree at most, and would be undetectable in aerial imagery. The largest one would be a significant fraction of all forest in region 8. Double-check these numbers.

We thank the reviewer for carefully examining these numbers. We agree that the polygon sizes reported in Table 1 may appear extreme at first glance.

Analyzing the raw IDS data for Region 8 before any filtering, we find that of the total 74,281 disturbances recorded, the largest was a snow-ice event covering 62,230.66 km², and the smallest was a southern pine beetle attack covering 0.00000041 km² (≈0.41 m²).

In the paper, we reported polygon sizes ranging from 0.5 m² to 62,231 km² (rounded), reflecting the dataset before the filtering applied to Region 8. These extreme values highlight the heterogeneous nature of the dataset: very small entries indeed likely represent single-tree disturbances or isolated points, whereas the largest events correspond to large aggregated polygons.

R1C7: The RQA TREND method for change detection is perfectly valid; however, the method has three parameters: the embedding dimension *m*, the recurrence threshold ε, and the delay τ. The results for the slope away from the main diagonal depends on them. None of the parameters is provided in the text; unfortunately, Cremer et al. (2020) does not mention them either; you also refer to the "European Commission…(2023)" proceedings, what you really mean is the Cremer et al. article on pp. 361-364 in that book (please be more precise in your referencing), but that article does not contain the values for the parameters either. Thus, the threshold for the trend -1.28 (l. 316 – what you probably mean is -1.28 yr-1) appears completely arbitrary (again, also this threshold is not mentioned in the Cremer et al. articles), what is its justification? It seems to be THE crucial parameter to distinguish non-disturbance to disturbance – how robust are your results against changing it? The extreme patchiness of the disturbance areas (e.g., as seen in Figure 7) could be a result of that choice. – You are also stating the opposite of what you intend to say in l. 316f, the correct version would be "Pixels with a RQA-Trend value ABOVE the threshold of -1.28 were considered to show no significant change." Please be more explicit here, and check the consequences of changing the trend threshold.

Thank you for the careful consideration of the RQA TREND Methodology. We agree that we should add more detail about these three parameters.

We have corrected the previously cited "European Commission…(2023)" proceedings to Cremer et al. (2020); as this paper represents the most recent and appropriate reference for this method.

Regarding the RQA parameters, following the framework implemented in Cremer et al. (2020), recurrence plots were constructed directly from the univariate Sentinel-1 backscatter time series without phase-space reconstruction. The RecurrenceAnalysis.jl package (the Julia implementation we are using, ***RecurrenceAnalysis.jl***) supports the use of embedded state vectors if provided, in this study, we did not perform time-delay embedding, and thus no embedding dimension or delay parameters were involved. The recurrence threshold ε was the sole parameter used to construct the recurrence matrix. As such, embedding dimension and delay parameters are not applied in this formulation; the recurrence condition is based solely on the absolute distance between scalar observations (Eq. 1) in Cremer et al. (2020). Consequently, the only tunable RQA parameter in our workflow is the recurrence threshold ε. The recurrence threshold (ε) was set to 3 dB. This choice has now been explicitly stated in the Methods section.

Unfortunately, the RQA TREND threshold was unfortunately not explicitly reported in Cremer et al. (2020), although this choice is justified in the original manuscript. The threshold used in the submitted version of the manuscript had been derived empirically from the 5th percentile of the TREND for a known bark-beetle disturbance event in Europe. The choice of 5th percentile is based on Kellndorfer (2019) as an appropriate value to detect forest loss in SAR data, and validated in Cremer et al. (2020) for case studies of deforestation. We agree that this choice is necessarily heuristic, even if it is grounded in observed data distributions rather than being arbitrarily selected, and that there might be uncertainties associated with the type of event and location (i.e. a bark-beetle event in Europe) used to derive the threshold value.

Therefore, to assess the robustness of this approach, we evaluated the effect of the threshold choice on the performance of the S1DM detection for a single event for each disturbance type (DCA_ID) by comparing Precision, Recall, F1 score and Jaccard similarity for varying values of RQA TREND (from -3.28 to 0.72, in steps of 0.5) in our reference labels based on Planetscope data (Figure R1.6).

The analysis shows that indeed, the choice of RQA TREND threshold has an important influence on the performance of the disturbance detection and delineation, as given by differences in the four metrics. Specifically, with increasing values of RQA TREND (i.e. from more negative to less negative), we find a decrease in precision and increase in recall for all three disturbances. For wind and bark-beetle, this results in a declining F1 score with increasing RQA TREND thresholds (i.e. worse performance for more sensitive detection), and we also find a degradation in the agreement of the delineation of the disturbances (Jaccard similarity). However, for defoliators we find an inverted parabolic shape for both F1 scores and Jaccard similarity, which peak at around -1.28 to -0.78, whereas stricter thresholds (less sensitive to small changes) and more sensitive thresholds both result in reduced performance.

The resulting best thresholds, based on the F1 scores and Jaccard similarity, per DCA_ID were: Bark Beetle −2.78, Defoliators -1.28 or −0.78 (marginally different), and Wind −2.78. These results demonstrate that the optimal threshold differs between disturbance types. To apply a single, average threshold applicable across all DCA_IDs, we retain our selected threshold of −1.28, which falls within the overlapping range of the optimal thresholds across the different disturbance types and evaluation metrics (Precision, Recall, F1 and Jaccard similarity). This choice ensures a reasonable compromise between detecting true disturbances and limiting overestimation. Our analysis shows that this average threshold generally performs well, and is particularly suitable to detect the more subtle nature of defoliator disturbances, while being able to capture wind and bark-beetle disturbances with reasonable skill.

Importantly, this analysis highlights that there is no universal “change-detection-fits-all” threshold as disturbance agents exhibit distinct spatial and temporal characteristics, leading to varying optimal detection thresholds. This is an important conclusion of this analysis which we consider to be worth highlighting, so that we now add it to the separate Discussion Section 5.3. *Methodological Assumptions and Limitations* and added Figure R1.6 to the Appendix of our manuscript, where it is numbered as Figure A.3.

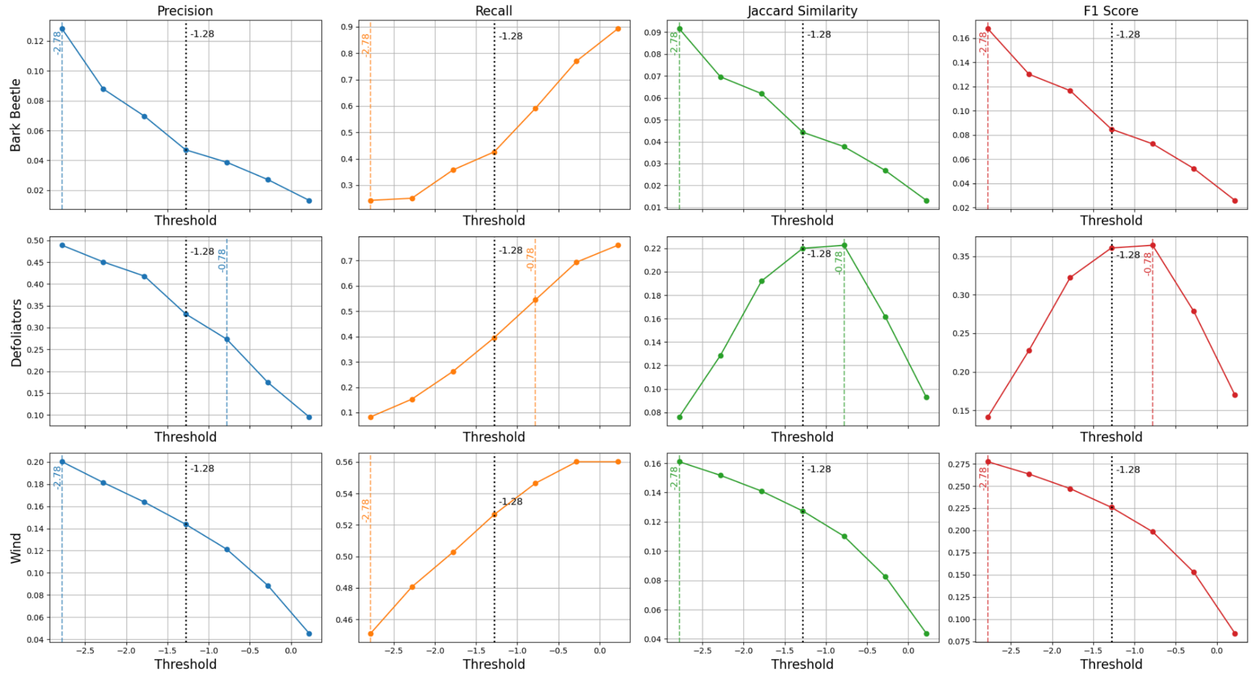

Figure R1.6: Threshold analysis for three events, one for Bark Beetle, Defoliators, and Wind each (top to bottom). Reference data are manually labeled disturbance patches derived from high-resolution Planet imagery, used to compare different Sentinel-1 change detection thresholds. Columns show the evaluated metrics. The x-axis represents the tested thresholds for RQA TREND. The black dotted line indicates the selected threshold of −1.28, while the colored dashed lines represent the F1-maximizing threshold for each disturbance type.

R1C8: Did you calculate one RP per time series, or did you move a window (e.g. of one year length) across the time series and calculated a separate RP and thus a TREND each time? If a disturbance sets in, it is to be expected that some of the RQA variables (TREND, but possibly also DET) "react" more or less suddenly (e.g., for wind). That would provide an opportunity to put a more precise date for the onset of the disturbance.

For the computation of RQA metrics, we calculated one recurrence plot and one TREND value per pixel using a fixed two-year time window centered within each year. Therefore, we use a moving window, even if this is done in annual steps. The reason for this choice is two-fold: first, our reference data (IDS) has an annual resolution, therefore, having sub-annual detail on the S1CD detection would not change the results, in the sense that we would still be performing the comparison on an annual basis; second, the RR algorithm is quite computationally expensive since it compares each time-step with all time-steps within the two year window, i.e. for 60 time steps in total, a total of 3600 points are calculated for the recurrence plot, per analysis year, per pixel. We agree that having a more refined moving window could yield interesting results in that it might allow to better quantify the seasonal timing of the disturbance, but given that this is not the main focus of this analysis, and that the bottleneck in terms of IDS temporal resolution would render that analysis superfluous for our goals and extremely costly in terms of computing time, we decide to keep our current approach.

The rewritten paragraph in the method section now states:

> L.297-331: "Based on the Sentinel-1 VH backscatter data for the years 2016–2021 over our study region described in Section 2.4, we applied the pixel-wise change detection method described by Cremer et al. (2020) to identify forest disturbance patches at 20 m x 20 m spatial resolution.
>
> The approach relies on Recurrence Quantification Analysis (RQA; Marwan et al., 2007) of time-series of VH data, which is applied to each pixel individually. The purpose of the RQA approach is to quantify the similarity between time steps in a time series. Here, for each calendar year within the study period (2016-2020), we select a two-year window spanning July of the previous year to June of the following year in order to evaluate the presence of a disturbance in that calendar year (i.e., 60 time-steps). The main elements of the RQA algorithm are illustrated in Fig. A2.
>
> In brief, the method plots the time series against itself to generate a Recurrence Plot (RP), where each point in the time series is compared with all others (i.e., 60 x 60 time-steps in this case), and a recurrence condition (true/false) is assigned based on the absolute difference of the backscatter signal. Here, the recurrence threshold (ε) was set to 3 dB. We computed the RP following Cremer et al. (2020) by using the RecurrenceAnalysis.jl package.
>
> By definition, the main diagonal is composed of ones (Fig. A2c). Then, a Recurrence Rate (RR) can be calculated for each diagonal of varying length, defined as the fraction of recurrent points along the diagonal relative to its length, resulting in a recurrence line quantifying the recurrence fraction per length (Fig. A2d). The slope of this line (RQA-TREND) reflects the temporal dynamics of the pixel: a slope of zero indicates no change over the whole period, whereas negative or positive slopes indicate decreasing or increasing recurrence, respectively. We applied an RQA-TREND threshold of −1.28 to distinguish between stable and disturbed pixels. This threshold was selected based on a preliminary analysis of the 5th percentile of the RQA-TREND over a known disturbance event, consistent with the results Cremer et al. (2020) for deforestation events. The choice of 5th percentile is based on Kellndorfer (2019) as an appropriate value to detect forest loss in SAR data, and validated in Cremer et al. (2020) for case studies of deforestation. To test the robustness and validity of this choice on our study area, we performed a sensitivity test on a subset of 3 Planetscope manually labeled disturbance patches (see Section 3.5), one for each disturbance (see Fig. A3).
>
> We tested a range of RQA-TREND thresholds and evaluated their performance using the F1 score and Jaccard similarity for each disturbance type (DCA_ID). The optimal

thresholds differed among disturbance agents, with stronger trends performing better in detecting bark beetles and wind (−2.78), and more moderate thresholds for defoliators (F1 peaking at -1.28 to −0.78, with very similar values).

Given the preliminary sensitivity test, we selected the threshold of −1.28, which lies within the overlapping range of effective thresholds across the three disturbance types. This compromise provides a balanced performance between detecting true disturbances and limiting overestimation, although it is somewhat restrictive for capturing defoliator disturbances. We classified pixels with RQA-TREND values above −1.28 as showing no significant change, and those below this threshold as indicating a structural change in that year. This process produced a binary raster per year, where each pixel indicated whether a structural change in forest cover occurred during that year. Finally, we projected this resulting WGS84 coordinate system (EPSG:4326) to ensure compatibility with the other spatial datasets used in this study.

To exclude non-forest areas from the raster, we used the preprocessed CONUS Tree Canopy Cover (TCC) dataset to filter out non-forest areas (refer to 3.2). Lastly, we converted the binary raster pixels into polygons to ensure a consistent, polygon-based dataset format.

In the following, we refer to this dataset as S1CD (Sentinel-1 Change Detection)."

Additionally, the new Discussion section **5.3. Methodological Assumptions and Limitations** now addresses these Limitations explicitly:

L650-668: [...] "Another key methodological limitation concerns the selection of the temporal threshold for classifying structural changes. Like most other disturbance detection approaches (van der Woude et al.,2026; Hirschmugl et al., 2017; Kennedy et al., 2010), our algorithm requires a change in the temporal structure of the VH signal that is sharp enough to be distinguished from normal variability. In this study, we applied a threshold of −1.28 of the RQA-TREND calculated over a 2-year window centered around each year's summer to distinguish between stable and disturbed pixels. While this value was empirically derived from reference datasets, it remains a heuristic choice and can influence the spatial coherence and apparent patchiness of detected disturbances.

We tested the effect of this choice by evaluating the performance of the S1CD in detecting the set of manually labeled disturbance patches from PlanetScope for different values of the RQA-TREND and per disturbance type (Fig. A3).

The results indicate that the optimal thresholds differ among disturbance agents, with wind and Bark-beetle requiring sharp changes in the temporal structure of the SAR signal, while defoliators are better captured if less strict thresholds are used.

The selected threshold of −1.28 falls within the broader range of optimal values across disturbance types and metrics, representing a compromise solution that ensures consistent applicability across all disturbance categories.

We explicitly acknowledge this dependency as a limitation of our workflow. The analysis highlights that, in principle, a universal "one-size-fits-all" change-detection threshold does not exist. These results are consistent with and underscore the challenges in monitoring from space disturbance with very different spatial and temporal characteristics, as discussed, e.g., by McDowell et al., (2015). We therefore recommend adopting disturbance-specific threshold calibration strategies, as global fixed thresholds are known to produce biased results across ecosystems and disturbance types due to differences in vegetation structure, disturbance dynamics, and sensor response (McDowell et al., 2015; Cohen et al., 2010; Verbesselt et al., 2010; Zhu and Woodcock, 2014; Senf and Seidl, 2018)." [...]

R1C9: Concerning the IDS dataset, you mention "over 1000 selectable agents" (l. 162). This is surely a source of uncertainty; how can any image interpreter ever pick the right one out of so many choices under time pressure etc.? How many of these 1000 did you have to aggregate to get to the broad categories "wind", "bark beetle", "defoliator"? Please discuss. What is the connection between the > 1000 choices and Table A.3 (the transformation of the choices into DCA_ID)?

Thank you for this question. We agree that the description of disturbance agents in the manuscript could lead to confusion and therefore warrants clarification.

In the IDS database, disturbance agents are organized into 28 broad groups (e.g., general insects, bark beetles, defoliators, chewing insects, general disease, biotic damage, domestic animals, and abiotic damage). Each group can be further subdivided into more specific subgroups or species-level categories, resulting in more than 1,000 distinct agent codes.

Disturbance agents are encoded using a five-digit numerical code, in which the first two digits identify the major disturbance group and the remaining digits specify subgroups or individual species. For example, bark beetles belong to the major group coded as *11***, where the generic bark beetle category is represented by code *11000* (bark beetles, *Scolytinae*). If surveyors identify a specific species, more detailed codes are used, such as *11007*, which represents for example the Douglas-fir beetle (*Dendroctonus pseudotsugae*). Despite this high taxonomic specificity, these agents remain within the same major group, bark beetles. Comparable hierarchical subdivisions exist for other disturbance categories, including abiotic subgroups such as drought or wind.

Also, note that the agent species are not identified solely based on the orthophotographs. Before each new flight season, Quick Key lists are prepared that summarize the disturbance agents, host species, and damage types most commonly expected within a given state, region,

or survey area, and these are provided to the surveyors. In addition, ground checks are conducted to support and validate aerial observations.
In the GIS Handbook of USDA [19], they write: “Regardless of survey method, ground checking should always complement data collection with real-time feedback to keep observers tuned and adjust future observations during operations. Ground checks should be done where core attributes are unknown, following any survey method. Ground checks should also be done where new damage types or agents are suspected, and always when regulated pests are suspected. The extent of ground checking is subjective but should be thorough enough that the observer is confident the ground checks are representative of the feature of interest in terms of spatial position and core attributes; adjustments to any features should be made prior to data submission.”

In our analysis, disturbance agents were aggregated to their highest hierarchical level (major group), except for wind, which (although formally a subgroup of abiotic damage) was treated as a separate category due to its distinct disturbance characteristics. We added Table 1.1 *(*numbered *Tab.* A.4 in the manuscript) to the Manuscript that displays the Category and number of subcategories within. Accordingly, the DCA_IDs in Table A.3 (of the manuscript) represent the broad categories (Bark Beetle, Defoliator, and Wind) rather than the detailed IDS agent codes.

We agree with the reviewer that surveyors' ability to distinguish disturbance agents at the species level represents a potential source of uncertainty. Therefore, the aggregation to the broader categories of bark-beetle, defoliator, and wind allows to reduce the uncertainty regarding the particular agent of each disturbed patch. We revised the manuscript accordingly.

> *L. 155-164: Disturbances are mapped as geo-referenced polygons that represent areas where damage was observed, though not all trees within a polygon are necessarily affected. For each polygon,* surveyors record a range of attributes **(USDA Forest Service, 2022, 2005); those used in this study include** the affected tree species*, the causal agent,* the survey year, *the damage type (e.g., mortality, defoliation percentage, crown dieback, branch breakage, or other non-mortality damage), and the severity class, expressed as the percentage of live and damaged trees within the polygon (Very Light: 1–3% through Very Severe: >50%).* ***For the causal agents, IDS database includes more than 1,000 fine-grained disturbance agent codes, which are hierarchically organized into 28 broader categories. To reduce interpreter uncertainty and ensure analytical consistency, the 63 bark beetle and 210 defoliator agents were aggregated to their broader categories in IDS (Table A.4), while wind, formally a subgroup of the abiotic damage category, was considered separately due to its distinct disturbance characteristics.. The resulting aggregated disturbance categories used here, namely wind, bark beetles and defoliators are referred to as DCA_ID.”***

Table 1.1: Codes of the 28 major disturbance-agent categories used in the IDS Damage Causal Agent (DCA) classification, along with the number of subcategories defined under each category and selected examples of subgroups, as described in the Aerial Survey Geographic Information System Handbook of the U.S. Forest Service (USDA Forest Service, 2005).

IDS Damage Causal Agent Codes	Major Disturbance Agent Category	Number of subgroups	Examples of Subgroups
10000	General Insects	16	Wasp, Ant, Ash Whitefly
11000	Bark Beetles	63	Spruce Beetle, Western Pine Beetle, Mountain Pine Beetle
12000	Defoliators	210	Bruce Spanworm, Aspen Leafminer, Western Spruce Budworm
13000	Chewing Insects	32	Grasshopper, Ash Plant Bug
14000	Sucking Insects	80	Spittlebug, Striped Mealybug
15000	Boring Insects	97	Termites, Roundheaded Borer
16000	Seed/Cone/Flower/Fruit Insects	56	Douglas-fir Cone Moth, Hollyhock Thrips
17000	Gallmaker Insects	23	Birch Budgall Mite, Spider Mites
18000	Insect Predators	5	Lacewings, Western Yellowjacket
19000	General Diseases	0	–
20000	Biotic Damage	4	Gray Mold, Damping-off
21000	Root/Butt Diseases	35	Armillaria Root Disease, Littleleaf Disease / Phytophthora Root
22000	Stem Decays/Cankers	90	Sap Rot, Viruses
23000	Parasitic/Epiphytic Plants	23	Mistletoe

24000	Decline Complexes/Dieback/Wilts	34	Yellow-Cedar Decline, Birch Dieback
25000	Foliage Diseases	79	Needlecast
26000	Stem Rusts	15	Pinyon Rust, White Pine Blister Rust
27000	Broom Rusts	6	Spruce Broom Rust, Juniper Broom Rust
28000	Terminal, Shoot, and Twig Insects	2	Pine Shoot Beetle
29000	Root Insects	0	–
30000	Fire	4	Wild Fire, Crown Fire Damage, Human-Caused Fi
41000	Wild Animals	19	Bears, Rabbits, Woodpeckers
42000	Domestic Animals	6	Domesticated Cattle, Sheep
50000	Abiotic Damage	22	Drought, Hail, Snow/Ice, Wind (Tornado/Hurricane)
60000	Competition	0	–
70000	Human Activities	15	Herbicides, Land Use Conversion, Roads, Harvest
80000	Multi-Damage (Insect/Disease)	4	Subalpine Fir Mortality, Five-Needle Pine Decline
90000	Unknown	11	Broken Top, Forked Top, Foliage Discoloration

R1C10: The annotated pdf attached to this review contains a further **31** comments, mostly rather specific ones. Please consider them as well.

We have carefully reviewed the annotated PDF supplement and have addressed each of these in line comments at the end of this response, providing detailed explanations and clarifications for each concern.

R1C11: The paper has some strong points on being self-critical, indicating the limitations of the study, the problems with spatial inaccuracy and thus the necessity to introduce a buffer zone around the polygons, etc. It becomes obvious that the three disturbance categories are very different in their spatial structure. Rendering IDS and S1DM truly comparable for bark beetle and defoliator attacks is a long way to go, as Figure 4 demonstrates.

We thank the reviewer for the kind comment. We have made every effort to be transparent about the limitations of our study, including the spatial inaccuracy of the data and the resulting need for buffer zones around polygons. We agree that the three disturbance categories differ substantially in their spatial structure, and as Figure 4 shows, fully rendering IDS and S1DM comparable for bark beetle and defoliator events remains a significant challenge. Our results demonstrate that Sentinel-1 offers a promising avenue for improvement, but this work represents only the first step toward fully integrating these datasets.

R1C12: A rather tricky issue is the timing of the onset of a disturbance; here, the authors go to a very coarse resolution of even more than one year, indicating that “online detection” of new damages is impossible. This is really a pity, since the strength of S-1 (and also S-2 for that matter) is short revisit cycles, with the potential to detect emerging attacks early and potentially act accordingly, very relevant for ecosystem managers! (See Jamali et al. 2023 for an approach using S-2). As the setup is now, the S1DM is for documentation of past events only.

We agree with the reviewer that this is a limitation of our current approach. While Sentinel-1 and Sentinel-2 offer short revisit cycles that could, in principle, enable early detection of emerging disturbances, our reference dataset (IDS) is only available at an annual resolution, and with an uncertainty of ca.±0.7 years (rounded to ~1 year) with a standard deviation of ±3.7 years (~4 years) compared to FIA, as discussed in Eifler et al. (2026). Furthermore, since the goal of our study is to develop a more refined reference dataset for training and benchmarking satellite-based disturbance models, it is focused on documenting past events. We note, nevertheless, that even with its annual resolution, S1DM tends to detect some disturbances earlier than IDS, particularly bark-beetle. Therefore, we hope our dataset will support the development of real-time or near-real-time detection algorithms in the future.

R1C13: The last sentence (l. 649f.) is talking about an “alternative to manual labeling”; ironically, you are judging the quality of the S1DM as compared to IDS based on a third dataset which was obtained by manual labeling. The “fully automatic forest disturbance classification methods” (l. 645) are still to be developed.

We acknowledge the reviewer’s critical perspective on these sentences. We would like to clarify that PlanetScope manual labels are not used as training data for the S1DM workflow. Instead, they serve as an independent quality-control dataset to assess whether our fully automatic method correctly captures disturbance outlines and timing correctly. Such quality control is necessary because the original IDS dataset contains spatial and temporal uncertainties (e.g., Eifler et al., 2026; Coleman et al., 2018), and although the Sentinel-1 method has been developed, published, and validated (Cremer et al., 2020), combining these two sources does not eliminate the need for independent verification. Therefore, manual labeling was used for validation, even though the S1DM pipeline is fully automated.

The changes suggested still add up to “minor revisions” only, this is regarding the text. The sensitivity analysis for the TREND threshold and the selection of PlanetScope/S1DM but NOT IDS damaged areas require additional work however.

Reference

Jamali, S., P.-O. Olsson, A. Ghorbanian and M. Müller (2023). "Examining the potential for early detection of spruce bark beetle attacks using multi-temporal Sentinel-2 and harvester data." ISPRS Journal of Photogrammetry and Remote Sensing 205: 352-366. https://doi.org/10.1016/j.isprsjprs.2023.10.013

Answering the **Minor comments** from **R1C10**:

L12: acronym not introduced before
Thank you for noticing. We have now introduced the acronym appropriately at first use.

L42: i.e. being reported?
Thank you, we have updated this sentence. It now reads:

> L. 43: “Coverage of reports on severe weather disturbances included 50% of forest areas in North America, 86% in Europe, and just 8% in Asia.”

L99: Remove "drought"
Thank you for noticing, this is a fragment of previous drafts and was removed alongside a large part of that paragraph.

Fig1: what about disturbances > 15 km2 in size? and where do the 15 km2 come from?
We addressed this concern in **R1C5**, where we explained our intentions for the 15 km² threshold. In this case, we keep the original Figure Caption but added a clarification in Section 3.1 IDS Preprocessing (see **R1C5**)**.** This addition makes our handling of extremely large disturbances explicit and justifies their exclusion from the analysis.

Please refer to **R1C5** for the revised text.

L162: how many of them are attributed correctly?

According to Coleman et al. (2018), the overall accuracy for identifying damage type was 97%, with a kappa statistic of 0.95. Their analysis was based on a subset of 567 aerial survey polygons from 2012 to 2014 that were ground-checked.

As explained in **R1C9**, we revised the sentence to clarify the mapping approach and simplified the description of disturbance agents, replacing "over 1,000 selectable agents" with "28 broader disturbance agents."

Please refer to **R1C9** for the updated sentence.

L166/Tab1: make sure you really mean km2 here, not m2. This largest polygon would cover almost 9% of the whole forested area of region 8.
The reviewer's comment overlaps with **R1C6**, which we have already addressed.

As also discussed in our response to **R1C5**, forest disturbances reported by IDS can occasionally span extremely large areas, in some cases covering several percent of the total forested region. As these very large polygons are likely attributable to human error or overly generalized mapping, we excluded them by restricting the analysis to disturbances smaller than 15 km².

L191: in time probably, not in space?
Exactly — in this context, *activity windows* refer to a temporal window, not a spatial one. It indicates the specific time frame during the year during which a biotic disturbance (e.g., insect defoliation or bark beetle attack) occurs. The revised sentence now reads clearly:

> L. 188-189: "Biotic disturbances, such as defoliators and bark beetles, occur within specific **temporal** activity windows that must be captured accurately through timely ADS."

L263: this is repetitive
Agreed. We have removed the sentence "Furthermore, our analysis focused on three specific types of disturbances: wind, bark beetles and defoliators.", as this point is explained later in the manuscript.

L269: deleted 'type'
Thank you. We agree and have deleted the word "type".

L278: that's drastic difference, what is the rational behind it?
We addressed this concern in **R1C5**, where we explained our decision on the 15 km² threshold, and updated the manuscript in Section 3.1, IDS Preprocessing. Please refer to **R1C5** for the revised text.

L295-296: since the pixels are small (20 m x 20 m), the 30% threshold might exclude quite a bit of low-density forest. According to FAO, the threshold is 10% (but over an area of 0.5 ha, or 12.5 S1 pixels) https://fra-data.fao.org/definitions/fra/2020/en/tad

We already addressed this concern in **R1C3**, where we reran the analysis using a less restrictive 10% Tree Canopy Cover (TCC) threshold as suggested by the reviewer. The results were very similar to the original analysis using 30% TCC, with only minor changes in the number of detected disturbance events and no change in the relative loss percentages. Given this, we decided to retain the 30% threshold for the main analysis to provide a more conservative and reliable estimate of forested areas.

Please refer to **R1C3** for the revised text.

L316. no, this has to be the magnitude (absolute value) of the RQA trend; i.e. all trend values in the intercal [-1.28; 1.28] are considered as no change. Where does the 1.28 come from? What was the embedding dimension used? What was the recurrence threshold \epsilon ? Need more details here

This comment has been addressed in **R1C7.**

L328-329: Why do you focus on these only?

The comment refers to the sentence in the manuscript: "*To focus on enhancing disturbance information rather than detecting new events, we only considered S1CD elements located within a buffer around IDS polygons, excluding all areas outside both the buffer and the flown survey paths.*"

We focused on these areas because the goal of this study is to refine the spatio-temporal attributes of existing disturbance datasets, rather than to detect new disturbances. Detecting new disturbances does not yet allow for reliable attribution of disturbance agents, which is ongoing work. Attributing disturbance types across all forest patches at a large scale is currently not feasible. Previous studies (Senf & Seidl, 2021; Viana-Soto & Senf, 2024) have focused on limited subsets of disturbance types, such as combined wind–bark beetle events, fire, and harvest. However, comprehensive disturbance attribution remains challenging—particularly when distinguishing among insect disturbance types (e.g., bark beetles and defoliators) and between insect- and wind-driven disturbances.

In our work, we focus on these disturbance types to enhance their spatial and temporal accuracy in legacy datasets, creating a high-quality reference that other researchers can later use to train classification algorithms and include for new events detected by remote sensing. By focusing on IDS-buffered areas, we ensure this enhanced dataset is reliable and suitable for future modeling beyond IDS-covered regions.

We revised the sentence in the manuscript.

L. 335-337: “*To focus on* ***refining the spatio-temporal attributes of known disturbances,*** *rather than detecting new events, we* ***considered only*** *S1CD elements located within a buffer around IDS polygons, excluding all areas outside both the buffer and the flown survey paths.”*

2 Comments: L387 & L391: Removed Jaccard similarity

We are unsure why the Jaccard similarity was suggested for removal, as it is an important metric for understanding our method. To improve clarity, we have revised the first mention to “Jaccard similarity index” and retained it throughout the manuscript.

L. 389-390: “For each polygon of IDS or S1DM, we computed the spatial overlap (Eq. 1) and the **Jaccard similarity index** (Jaccard, 1902) (Eq. 2) with the corresponding manual polygon.”

Figure5: a)1: any explanation for this peak in the pdf for both IDS and S1DM at around 5 km2 ?
We thank the reviewer for this important observation.
An analysis of IDS data without removing compound events shows that the peak in the probability density function around 4-6 km² for wind disturbances in 2020 is associated with unusually large events, primarily concentrated in a specific region of Louisiana, between Houston and New Orleans. Many of the polygons in this region are square-shaped, suggesting that part of the observed pattern may reflect mapping artifacts rather than the exact spatial extent of disturbances.

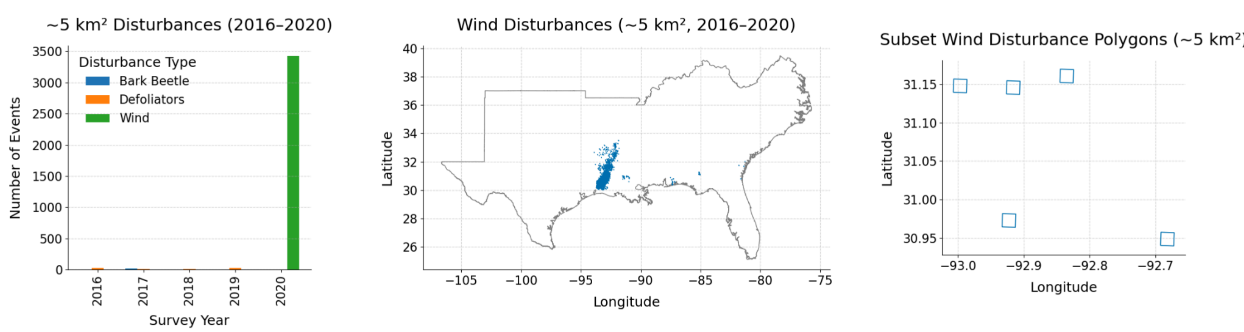

Figure R1.7: Left: Bar plot showing the number of disturbance patches with areas between 4–6 km², grouped by disturbance type (bark beetles, defoliators, wind) and survey year. Bar colors indicate disturbance type, and the y-axis represents the number of disturbance polygons. Center: Spatial distribution of all wind disturbance polygons with areas between 4–6 km² detected between 2016 and 2020 across the study region. Right: Detailed view of a subset of wind disturbance polygons (4–6 km²) detected in 2020, with individual polygon boundaries (blue) shown to illustrate disturbance shape.

There are two main factors contributing to this pattern:

1. Actual large disturbances: In 2020, the region experienced major wind damage from several tropical cyclones, including Tropical Storm Cristobal, Hurricane Marco, Hurricane Laura, Hurricane Delta, and Hurricane Zeta. In particular, Hurricane Delta caused extensive forest damage, producing naturally large disturbance areas that contributed to the large wind-driven forest destruction in the data IDS and S1DM.

2. Mapping artifacts due to COVID-19 restrictions: The USDA modified its aerial survey protocols in 2020 because of pandemic-related restrictions. Standard aerial flights were reduced or not conducted, and alternative methods, including ground checks and imagery interpretation, were used instead. As a result, many IDS polygons in this region are square-shaped and approximately 5 km² in area, reflecting changes in survey methodology rather than actual disturbance extents.

Together, these factors explain the peak at ~5 km² in Figure 5: it arises from both true large-scale wind disturbances and methodological artifacts arising from COVID-affected survey procedures, producing unusually uniform polygon sizes while still capturing the impact of these destructive storms.

We have added the following information in the results and discussion sections of the manuscript.

> L. 435-437 (Results): **“***Figure 6a shows a pronounced peak in the probability density function at approximately 4–6 km² for wind-related disturbances, which is consistently observed in both the IDS and S1DM datasets.”*
>
> L. 540-546 (Discussion): “The peak at approximately 5 km² for wind disturbances reflects a combination of genuine large-scale wind damage and methodological artifacts in the IDS data. In 2020, several major tropical cyclones (including Cristobal, Marco, Laura, Delta, and Zeta) affected the region between Houston and New Orleans, producing extensive areas of wind-disturbed forests. In addition, aerial survey procedures were modified due to COVID-19 restrictions, leading to the use of alternative mapping approaches. This resulted in more uniform, often square-shaped disturbance polygons of similar size, particularly in this region, which contributes to the observed peak. Despite this artifact, the mapped polygons still capture the spatial footprint of severe wind-induced forest disturbances.”

12 Comments in the following paragraph (**L431–436**):
Line 434: Removed: 50% of S1DM Wind events being 0.1 km2 and 90% being 0.03 km2 smaller than IDS.
434-435: not “unlike” if numbers in Table 2 are correct
L435-436: Deleted the red underlaid parts in the following sentence: For Defoliators, more than 50% of S1DM events are 0.01 km2 larger than their IDS counterparts, and 90% of events are nearly twice as large in S1DM as in IDS.

Thank you for your careful review and for catching this inconsistency. You are absolutely correct: the values reported in Table 2 are correct, but the accompanying text had not been updated from a previous version. We have therefore completely reworded the paragraph (previous manuscript L.431–436) to ensure full consistency with the table. The revised text now reads:

> L.432-439: “Overall, disturbance areas derived from S1DM closely match those from IDS, with only minor differences across disturbance types. Bark beetle disturbances show the strongest agreement, with nearly identical median areas (IDS: 0.0046 km^2; S1DM: 0.0043 km^2) and 90% of events in both datasets smaller than 0.08 km^2.
>
> For wind disturbances, median S1DM events are slightly larger than IDS (0.34 vs. 0.24 km^2), while the 90th percentile areas are nearly identical (5.03 vs. 5.06 km^2). Figure 6a shows a pronounced peak in the probability density function at approximately 4–6 km^2 for wind-related disturbances, which is consistently observed in both the IDS and S1DM datasets.
>
> Defoliator disturbances tend to be larger in S1DM than in IDS, with similar median values (0.50 vs. 0.49 km^2) but substantially larger upper-tail events, where the 90th percentile in S1DM is nearly twice that of IDS (6.33 vs. 3.68 km^2).”

L466: Deleted majority in sentence: “While a majority of events — 359 out of 899 (39.9 %) — are detected by S1DM in the same year as IDS (Lag=0), a substantial number fall within the ± 1 ±1- year window around the IDS detection year, with 140 events (15.6 %) detected one year earlier and 157 events (17.5 %) one year later.”

Absolutely correct! 39% is not a majority; rather, it is the largest proportion of events. We have revised the sentence to now read:

> L. 469-471: “While **the largest proportion** of events, 359 out of 899 (39.9%), are detected by S1DM in the same year as IDS (Lag = 0), a substantial number fall within the ±1-year window around the IDS detection year, with 140 events (15.6%) detected one year earlier and 157 events (17.5%) one year later.”

L518: That number is not provided in the reference

The original sentence read: “In this case, the low sensitivity of S1CD to subtle and slow-onset disturbances might stem from the change detection threshold, set at -1.28 (1/yr) by European Commission. Joint Research Centre. (2023).”

This is correct, the exact threshold was not reported in the reference. We updated the citation to Cremer et al. (2020). The threshold itself was decided by us based on the method described in this paper and in correspondence with the authors. Accordingly, the sentence has been revised to:

> L. 517-519: “In this case, the low sensitivity of S1CD to subtle and slow-onset disturbances may stem from the change detection threshold, which we set at −1.28 based on the method described in Cremer et al. (2020) and correspondence with the authors.”

L529: Three?

Yes, thank you for noticing. Indeed we meant three disturbances, and the manuscript has been corrected accordingly.

Hybrid forest disturbance classification using Sentinel-1 and inventory data: a case-study for Southeastern USA

Franziska Müller, Laura Eifler, Felix Cremer, Pieter Beck, Gustau Camps-Valls, and Ana Bastos, *EGUsphere*

Response to Reviewer #2

Referee Comment 2

Hybrid forest disturbance classification using Sentinel-1 and inventory data: a case study for the Southeastern USA

General comment:

This manuscript combines Sentinel-1 with IDS data for disturbance mapping (wind, bark beetles, defoliators) in a forest region, addressing legacy dataset limitations. However, **the text structure is often confusing and lacks explicit research questions.** Pervasive formatting inconsistencies further impede readability. Minor revisions for structural clarity, **quantitative validation beyond manual polygon delineation using PlanetScope**, and thorough clean-up are essential before resubmission.

We will revise the manuscript for clarity and readability, and we will revise the motivation statement in the introduction.

Specific comments:

R2C1: The introduction is well explained but provides excessive detail on general topics such as DL training, labeling challenges, and legacy datasets, which are not clearly linked to the study's focus on bark beetle, defoliator, and wind disturbances. The research questions or specific objectives the authors intend to address should be stated explicitly.

We thank the reviewer for this comment. While the Introduction provides important context, we acknowledge that some sections contain general information not directly linked to our study focus. To clarify, we specifically focus on wind, bark beetle, and defoliator disturbances because these disturbance types are particularly challenging to attribute and differentiate in large-scale datasets. For example, in the European Forest Disturbance Map (Senf and Seidl, 2021), only wind and fire disturbances are clearly distinguished, and in the European Forest Disturbance Atlas (EFDA; Viana Soto and Senf, 2024), fire and harvest are emphasized, while the "wind–bark beetle" complex cannot be disentangled. Insects, in particular any insects apart from bark beetle outbreaks, are often underrepresented and bark beetle outbreaks are difficult to separate from wind disturbances. Therefore, our study targets these specific disturbance types to address this critical gap and improve understanding of their dynamics and detectability.

To address this comment, we have rewritten the entire paragraph to clarify the motivation for the selected disturbance types and clearly state our central research question, which now reads:

> *L.94-111: "In this study, we develop a hybrid forest disturbance benchmark dataset that integrates satellite observations with disturbance inventory data, focusing on three key disturbance types: bark beetle, defoliators, and wind. This choice is motivated both by the known challenges and persistent gaps in remote sensing–based classification of these disturbance types (McDowell et al., 2015; Kautz et al., 2017; Schleeweis et al.,2020) and specifically in separating wind and biotic disturbances (Schleeweis et al., 2020, Viana-Soto & Senf, 2025), as well as their ecological importance (Seild et al., 2017; Hicke et al., 2020; USDA Forest Service, 2015; Heaton et al., 2023). For example, bark beetles have killed approximately 3.8 billion trees in western North America between 1997 and 2018 (Hicke et al., 2020), and the defoliating gypsy moth (Lymantria dispar) has affected over 21.5 million ha in the northeastern United States from 1924 to 2015 (USDA Forest Service, 2015). Insect disturbances are responsible for substantial carbon losses (10 ± 1.3 Tg C yr^{-1}), comparable to those attributed to fire (7 ± 1.0 Tg C yr^{-1}) in the United States (Harris et al., 2016). Wind disturbances, while sparse in time, can be catastrophic, as exemplified by Hurricane Hugo, which damaged around 1.8 million ha of forest in South Carolina (Heaton et al., 2023), and storms have been associated with carbon losses of up to 5 ± 0.7 Tg C yr^{-1} during the period 2006–2010 (Harris et al., 2016). Here we aim to support advances in wind and insect disturbance classification by combining C-band Synthetic Aperture Radar (SAR) data from Sentinel-1 with inventory-based information. As the SAR signal is sensitive to structural and moisture changes that are not easily captured by optical sensors and visual inspection and can penetrate through clouds, and given Sentinel-1's spatially and temporally continuous information, this approach allows to refine inventory based disturbance maps and reduce spatial and temporal uncertainties. To support validation and quality control, we additionally incorporate a Tree Canopy Cover (TCC) dataset, which provides independent high-resolution estimates of forest canopy cover, and use high-resolution Planet imagery with manual labeling to assess whether this approach leads to significantly improved disturbance detection."*

R2C2: [Results] Despite the preprint relying on spatial agreement and manual labels without such metrics, I missed a confusion matrix analysis quantifying omission and commission errors for each disturbance type (wind, bark beetles, defoliators). This quantitative error breakdown is essential for transparent S1DM validation and comparability with IDS/alternatives.

We thank the reviewer for this constructive comment and agree that additional information on omission and commission errors is valuable for understanding the capabilities of Sentinel-1 Change Detection (S1DM) to improve legacy inventory datasets.

Since S1DM is based on overlapping patches with IDS, we are only able to show omission errors - that is, cases where S1DM fails to detect disturbances identified in IDS. Commission errors cannot be directly assessed for several reasons:

1) First, the IDS data does not provide full spatial coverage of the USA territory, but rather of selected areas that are flown-over each year. This means that there might be areas correctly identified by S1CD as disturbed that do not show overlap with IDS because the area has not been flown over (see a more in-depth discussion of the uncertainties of IDS in Eifler et al. (2026)).
2) Second, this means that it is not possible to provide an estimate of commission errors per disturbance type.
3) Finally, the S1CD signal also includes signals from other disturbance types that are not those considered here (fire, drought, other abiotic damage, other biotic damage, harvest), so that a comparison of the aggregated disturbances detected by S1CD and IDS would be meaningless.

To address this, we added a new figure Figure 2.1 (Manuscript Figure 4) showing the percentage of IDS events that overlap with S1DM, highlighting omission errors, and included a corresponding paragraph in the Results section.

> L. 411-415: "*Figure 3 and Fig. 4 show that the detection effectiveness of S1DM varies by disturbance type. Wind disturbances have the highest agreement, with 98.1 % of IDS wind events having a corresponding S1DM signal within the 500 m buffer (469 of 478 events). Bark beetle disturbances follow, with 76.3 % of IDS patches matched by S1DM (899 of 1,177 events). Agreement is lower for defoliator disturbances, where only 67.6 % of IDS disturbed areas are captured by S1DM (144 of 213 events). The percentages in Fig. 4 refer to totals across all years and match the absolute numbers reported in Fig. 3.*"

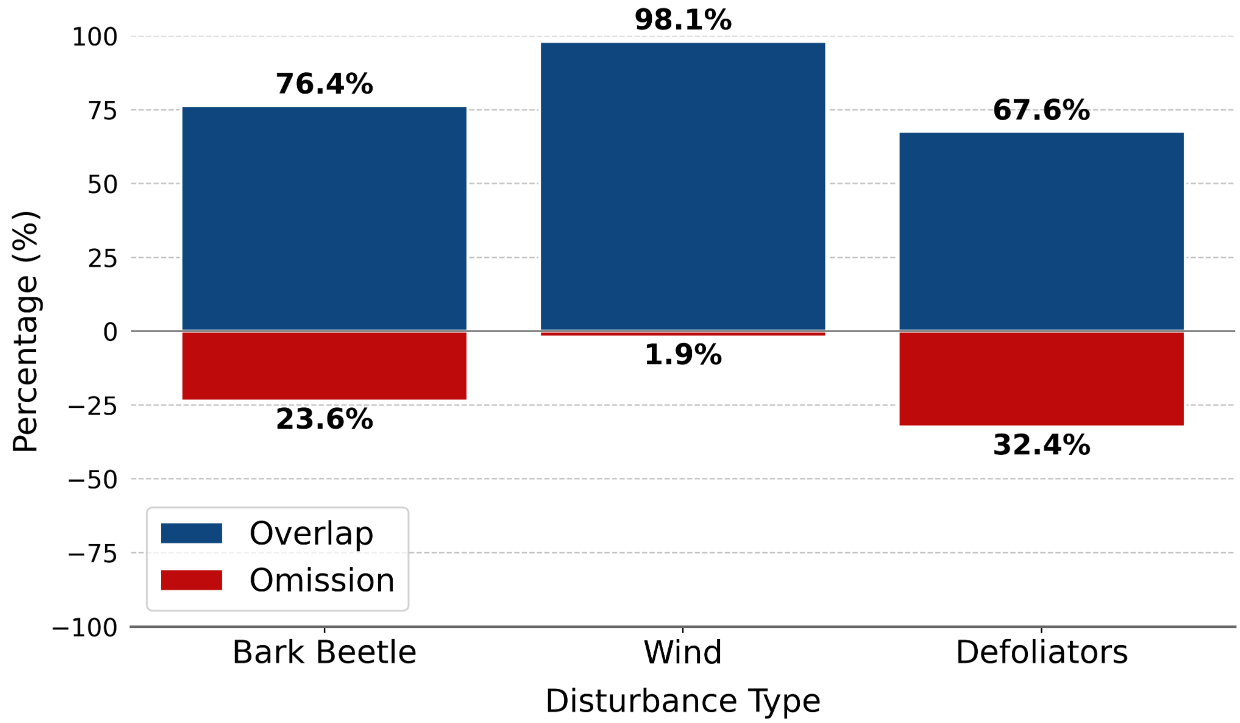

Figure 2.1: Overlap (blue) and omission (red) of IDS disturbance events with Sentinel-1 detections. The x-axis shows disturbance types (Bark Beetle, Wind, Defoliator), while the y-axis indicates the percentage of IDS events that were either detected by the Sentinel-1 Change detection or omitted. Blue bars represent IDS events with overlapping Sentinel-1 detections, and red bars show IDS events without a corresponding Sentinel-1 detection (omission).

[Discussion:] The section effectively highlights S1DM strengths but omits essential components: a dedicated limitations subsection (buffer choices, IDS error propagation, lack of field validation, regional limits); quantitative benchmarking against alternatives such as GLAD/ESDAC; explicit linkage back to the research questions with key results summaries; and implications for operational scalability and DL training. These omissions weaken the broader impact and dataset positioning.

We thank the reviewer for these constructive suggestions, which helped strengthen the Discussion section. To address these points, we have made the following decisions and revisions:

R2C3: A dedicated limitations subsection (buffer choices, IDS error propagation, lack of field validation, regional limits)

We would like to note that the implications of buffer choices were already discussed in Section 5.1, *Capabilities and Limitations of Sentinel-1 Change Detection to Improve Forest Disturbance Monitoring*, specifically in Section 5.1.1 *Location and delineation of disturbance patches*.

In response to the reviewer's comment, we have now added a dedicated Section 5.3 *Methodological Assumptions and Limitations*, which consolidates and expands the discussion of limitations across the workflow. This new subsection explicitly addresses buffer selection, error propagation within the IDS framework, and the lack of direct field-based validation. It reads:

> L. 633-690: "5.3. *Methodological Assumptions and Limitations*
>
> *In this study, several methodological decisions had to be made that could potentially influence our results.*
>
> *To combine disturbance information from IDS and S1CD, we used a spatial buffer of 500 m and a temporal buffer of ±2 years centered on the IDS disturbance date. This spatial and temporal buffer was introduced to account for the spatial and temporal uncertainty in both the IDS dataset (Eifler et al., 2026) as well as in the disturbance date detection by S1CD. Previous work by (Eifler et al., 2026) analyzed the spatial and temporal agreement in reported disturbances for IDS and the Forest Inventory and Analysis (Woudenberg et al., 2010) for the continental U.S. An analysis of spatial buffers of 800 m and 1600 m, showed that increasing buffer size introduced additional uncertainty. The disturbance timing differed by 0.7 years on average, for a buffer size of 800 m.*
>
> *We have tested the effect of using different spatial buffer sizes of 100 m, 250 m, 500 m, and 1000 m on the number of events retained for analysis, and found that buffer sizes of 100-250 m resulted in too few events being selected, while 1000 m did not retain considerably more events.*
>
> *Furthermore, our change detection algorithm is based on recurrence analysis trends over 2-year-long windows, meaning that the difference in the timing of disturbance with IDS could be up to 2 years, if the disturbance happened at the end of the analysis period. Therefore, our choice of 500 m and 2-year windows represents a balance between retaining information and minimizing the risk of introducing additional uncertainties. Future work could explore in greater detail the impact of spatial uncertainty between inventory datasets and satellite observations to refine buffer selection.*
>
> *Another key methodological limitation concerns the selection of the temporal threshold for classifying structural changes. Like most other disturbance detection approaches (van der Woude et al., 2026; Hirschmugl et al., 2017; Kennedy et al., 2010), our algorithm requires a change in the temporal structure of the VH signal that is sharp enough to be distinguished from normal variability. In this study, we applied a threshold of −1.28 of the RQA-TREND calculated over a 2-year window centered around each year's summer to distinguish between stable and disturbed pixels. While this value was empirically derived from reference datasets, it remains a heuristic choice and can influence the spatial coherence and apparent patchiness of detected disturbances.*

We tested the effect of this choice by evaluating the performance of the S1CD in detecting the set of manually labeled disturbance patches from PlanetScope for different values of the RQA-TREND and per disturbance type (Fig. A3).

The results indicate that the optimal thresholds differ among disturbance agents, with wind and Bark-beetle requiring sharp changes in the temporal structure of the SAR signal, while defoliators are better captured if less strict thresholds are used.

The selected threshold of −1.28 falls within the broader range of optimal values across disturbance types and metrics, representing a compromise solution that ensures consistent applicability across all disturbance categories.

We explicitly acknowledge this dependency as a limitation of our workflow. The analysis highlights that, in principle, a universal "one-size-fits-all" change-detection threshold does not exist. These results are consistent with and underscore the challenges in monitoring from space disturbance with very different spatial and temporal characteristics, as discussed, e.g., by McDowell et al. (2015). We therefore recommend adopting disturbance-specific threshold calibration strategies, as global fixed thresholds are known to produce biased results across ecosystems and disturbance types due to differences in vegetation structure, disturbance dynamics, and sensor response (McDowell et al., 2015; Cohen et al., 2010; Verbesselt et al., 2010; Zhu and Woodcock, 2014; Senf and Seidl, 2018).

A key limitation in disturbance detection workflows is the propagation of errors through sequential processing steps; for example, a misclassification in one dataset can influence the final combined disturbance map. Our framework is explicitly designed to reduce error propagation by integrating multiple sources of disturbance information, including Sentinel-1 change detection. While direct field validation would provide the strongest confirmation of disturbances, this is not possible in our case: the events we analyze occurred 5–9 years ago over a very large region, and many affected areas have likely recovered (for example, defoliator impacts often recover within a year). Previous work has explicitly assessed inconsistencies and uncertainties in IDS. Eifler et al. (2026) compared IDS with independent inventory data and found good temporal agreement with FIA, suggesting that temporal errors in IDS are generally low. Our results support this finding, as IDS timing largely aligns with S1DM (Fig. 7), except that some bark beetle disturbances are detected earlier by S1DM than reported in IDS. However, because FIA lacks sufficiently detailed spatial information, spatial event-level errors cannot be quantified using inventory data alone. To address this limitation, we employed Planet imagery as an independent, spatially explicit proxy for validation. Although this cannot fully replace field observations, it allows us to evaluate the plausibility of disturbances in areas where ground data are no longer available.

In this study, our analysis was constrained to areas covered by the IDS dataset, which provides detailed, spatially explicit information on disturbance agents. While Sentinel-1 offers global coverage and the potential to detect disturbances beyond regions surveyed by aerial or ground-based inventories, the lack of independent reference data limits our ability to confidently attribute detected disturbances to specific causal agents in areas outside IDS coverage. Consequently, the current framework cannot fully evaluate the performance of Sentinel-1-based disturbance detection across unsurveyed regions or in landscapes where IDS data are absent. Future work should address these limitations by integrating Sentinel-1 disturbance detection with high-resolution, independent reference datasets, such as manual delineations from Planet imagery, and by using tree canopy cover products to exclude non-forested areas. Expanding the analysis in this way would allow for a more comprehensive assessment of Sentinel-1's capabilities and its potential for large-scale, autonomous forest disturbance monitoring."

R2C4: Quantitative benchmarking against alternatives such as GLAD/ESDAC.

Thank you for pointing this out. After further investigation, we regret to report that it is not possible to incorporate these datasets into our analysis as a comparable quantitative benchmark.

The GLAD Forest Alerts dataset (https://glad.umd.edu/dataset/glad-forest-alerts) consists of two alert systems designed to detect tree cover loss. However, both systems fall outside the scope of our study region. GLAD-L provides tree cover loss alerts only between 30°N and 30°S, while GLAD-S2 focuses exclusively on primary humid tropical forests, primarily within the Amazon Basin. As our study area is located in the southeastern United States, neither product provides spatial coverage suitable for direct comparison. However, Eifler et al. (2026) evaluated the consistency of forest disturbance datasets across the continental United States and compared IDS with satellite-based products, including the North American Forest Dynamics (NAFD) and Global Forest Change (GFC) datasets, the latter based on a similar Landsat time-series approach. They found substantial temporal discrepancies between inventory- and satellite-derived disturbances. Specifically, the mean temporal lag of IDS relative to satellite products was approximately 0.5 years compared to GFC (±3.7 years) and 1.9 years compared to NAFD (±3.2 years), highlighting considerable disagreement in disturbance timing between inventory and remote sensing datasets.

With ESDAC, we are not sure if the reviewer was referring to the European Soil Data Centre, which does not directly report forest loss or forest disturbance data. Instead, it serves as a key provider of soil and land-related datasets, often used to analyze or contextualize forest loss processes. Upon reviewing the datasets, we understand that ESDAC primarily hosts datasets with European geographic coverage and does not cover our study region. Furthermore, since ESDAC does not provide forest loss or disturbance products per se, it cannot serve as a quantitative benchmark for our analysis.

Given these geographic and thematic limitations, neither GLAD nor ESDAC datasets can be used as a suitable comparison for our study, which focuses on forest loss dynamics in the southeastern United States. Alternative datasets in the USA (our study area) exhibit substantial uncertainty, whether derived from inventories or remote-sensing classifications (Eifler et al., 2026). Studies such as Andrus et al. (2025) further demonstrate that bark beetle-driven disturbances exhibit significant spatiotemporal variability (duration, cumulative mortality area (severity), maximum annual mortality, and mortality rate) that is difficult to capture with broad-scale remote sensing products alone. Together, these findings highlight the ongoing need for improved delineation methods that can accurately capture both the extent and the agent of forest disturbances.

R2C5: Explicit linkage back to the research questions with key results summaries

We agree with the reviewer that adding a specific ink back to the research questions will strengthen the discussion. So we reworded the beginning of the discussion:

> L.502-507: "*We aimed to improve forest disturbance mapping by combining IDS inventory data with Sentinel-1 radar change detection, leveraging IDS's detailed disturbance-agent information while reducing uncertainty in disturbance size, location, and timing. We find that integrating radar data substantially improves the spatial and temporal characterization of disturbance events, particularly for bark beetle outbreaks and wind damage. The resulting dataset captures both the location and timing of disturbances more accurately than IDS alone. These improvements provide a robust foundation for applications such as training data-intensive, data-driven models for disturbance classification."*

R2C6: Implications for operational scalability and DL training

We added a dedicated Section 5.4 *Generalizability and applications of the proposed approach*, which discusses the implications and benefits of this dataset.

> *L. 691-737: "Generalizability and applications of the proposed approach*
>
> *Our study demonstrates that integrating Sentinel-1 spatio-temporal information can help refine disturbance mapping from inventory datasets. The data processing pipeline is provided alongside this study (see GitHub in Section A), so that it can be extended to other regions, as well as different combinations of inventory and satellite-based datasets. We discuss some possible applications and considerations regarding the generalizability of this approach below.*
>
> *The approach proposed here can be easily applied to other inventory datasets and satellite missions, provided they have overlapping spatial and temporal coverage. The*

structural information inherent in SAR backscatter data and the high spatial resolution of Sentinel-1 offer opportunities to enrich existing forest disturbance inventories, improve temporal and spatial resolution, and reduce uncertainty in forest loss mapping. For IDS, which is already spatially explicit, S1CD improved the spatial delineation of the disturbed patches, especially for bark beetles and wind (Fig. 9). Therefore, our study shows that SAR-based change detection provides a valuable complementary source of information to aerial surveys and field inventories. For point-based inventory datasets, such as National Forest Inventories (Woudenberg et al., 2010), S1CD provides a means to expand the spatial information of the data, provided that precise coordinates are available for forest inventory plots. The pipeline can also be adapted to other existing satellite-based forest disturbance datasets, such as the Global Forest Change (Hansen et al., 2024), or the European Forest Disturbance Atlas (Viana-Soto and Senf, 2025). The main bottleneck for extending the application to other regions remains the limited availability of suitable inventory datasets, i.e., with large-scale spatial and temporal coverage and detailed and consistent forest disturbance agent information, ideally spatially explicit and with precise coordinates, and that are publicly available and that follow FAIR principles (findable, accessible, interoperable, and reusable). However, recent efforts to compile information on specific forest disturbances, (e.g. Forzieri et al., 2020, 2023; Urquiza-Muñoz et al., 2024) might still be useful. With appropriate reference datasets, however, there are no inherent technical limitations that prevent extending the methodology to other regions or larger spatial extents. Therefore, initiatives such as European Union's proposal for a Forest Monitoring Framework (European Commission, 2023), aiming to facilitate access to forest-relevant data, including forest disturbances (Migliavacca et al., 2025) would allow expanding and testing our approach in other regions.

An open question that remains is whether this approach is transferable to regions with complex topography, such as mountainous areas. Our study region is limited to the southeastern United States, which is relatively flat and was selected because of its high incidence of wind disturbances. The absence of large mountains avoids topographic-induced uncertainties known to affect Sentinel-1 observations (Shi et al., 2024; Borlaf-Mena et al., 2020). While the framework performs well under these conditions, its application in more mountainous regions would require additional testing. In addition, differences in forest structure and species composition may require region- or forest-type–specific adjustments to parameters, particularly for backscatter dynamics and disturbance signatures. In this context, other satellite-based information sources could replace, or complement the S1CD mapping, for example, vegetation indices reflecting other canopy properties from spectral imagery (Senf et al., 2015; Senf and Seidl, 2018; Canadell et al., 2021b; Hall et al., 2016), or hyperspectral information from new missions such as ENMAP (Vanguri et al., 2024).

Various forest disturbance classification algorithms have been developed to separate different disturbance types (Schleeweis et al., 2020; Senf and Seidl, 2021b; Senf et al., 2015; Oeser et al., 2017; Helmer et al., 2010; Baumann et al., 2014). The three disturbances considered here, wind, bark beetles, and defoliators, are particularly hard *to disentangle, given their particular spatial and temporal characteristics (e.g. McDowell et al., 2015; Schleeweis et al., 2020; Viana-Soto and Senf, 2025). With increasing interest in exploring machine learning (ML) or deep learning (DL) models for forest disturbance classification (Stahl et al., 2023; Kislov et al., 2021; Chen et al., 2021), having large-scale spatially and temporally consistent data on the different disturbance types is key either as training labels for supervised models, or else as evaluation data for new classification models in general. The S1DM dataset addresses this need, in that it contains a subset of IDS-based disturbed patches for these three disturbance types, with refined spatial delineation (Fig. 9) and temporal information, as S1CD tends to detect bark beetle, and to some extent also defoliator, disturbances earlier than they are reported in IDS (Fig. 7). This reduced set of disturbed patches is still large enough (469, 899, and 144 samples for wind, bark beetles, and defoliators, respectively, over a region of 1701214.39 km²) to support the development of dedicated and data-intensive classification algorithms to separate these different disturbance types. In this way, the refined dataset provides a foundation for the development of adaptive classification frameworks that can generalize across regions, forest types, and disturbance regimes."*

[Conclusion:] The conclusion requires substantial restructuring to enhance its impact and avoid redundancy. The first paragraph repetitively restates the global data gap and IDS limitations already discussed in the Introduction and Discussion. While subsequent paragraphs adequately summarize the hybrid S1DM method, they underemphasize the ecological implications—for example, how refined timing enables bark beetle outbreak forecasting, realistic wind patches inform gap dynamics and succession, or defoliator improvements correct carbon flux biases. The authors should explicitly position S1DM as a public benchmark for DL training (linking back to the labeling challenges in the Introduction) and replace vague future work with specific proposals.

We agree with the reviewer and thank them for helping us make a stronger conclusion. We have addressed the individual comments below and added the new conclusion, with the changes in bold, at the end.

R2C7: Shorten or remove the first paragraph repetitively restates the global data gap and IDS limitations already discussed in the Introduction and Discussion.

We thank the reviewer for this suggestion. We have shortened the first two paragraphs of the Conclusion to avoid repeating points already made in the Introduction and Discussion.

R2C8: Emphasize the ecological implications- for example, how refined timing enables bark beetle outbreak forecasting, realistic wind patches inform gap dynamics and succession, or defoliator improvements correct carbon flux biases in the conclusion.

We have added a discussion of ecological implications, highlighting that refined disturbance timing and spatial resolution can support improved modeling of forest dynamics, such as forest disturbance detection and classification, bark beetle outbreak forecasting, and correction of carbon flux changes caused by disturbances.

R2C9: Link the DL and ML from the Introduction and explicitly position S1DM as a public benchmark for DL training.

We have strengthened the link to the Introduction and explicitly positioned S1DM as a public benchmark for future forest classification algorithms, including ML and DL model development and evaluation.

The new conclusion (bold text indicates changes) with all suggestion from R2C7-R2C9:

> L.739-774: "***Reliable, spatially and temporally consistent information on forest disturbances, particularly wind, bark beetles and defoliator insects, remains a major bottleneck for large-scale disturbance attribution and modeling (Kautz et al., 2017, Rodríguez Paulino et al., 2024). In this study, we directly address this limitation by refining the widely used legacy Insect & Disease Survey (IDS) through the integration of Sentinel-1–based radar change detection (S1DM). This hybrid approach leverages the complementary strengths of both datasets: the long-term, agent-specific information provided by IDS (U.S. Forest Service, 2024) and the spatially and temporally continuous, cloud-independent structural sensitivity of C-band SAR observations.***
>
> *Our analysis specifically targets three key sources of uncertainty in legacy disturbance inventories: disturbance location, disturbance outline, and disturbance timing* ***(Eifler et al., 2026; Coleman et al., 2018; Andrus et al., 2025)****.*
>
> *Although S1CD shows reduced sensitivity to subtle canopy changes, such as those associated with defoliator activity, it performs well for structurally significant disturbances, including wind damage and bark beetle outbreaks****, a performance consistent with Bruggisser et al. (2021), that demonstrated that Sentinel-1 can reliably detect decreases in forest height greater than 10 m.*** *Spatially, disturbance centers were generally within 200–950 meters of one another. Although the overall disturbed area remained similar, the S1DM dataset identified more physically plausible disturbance patterns, with Wind and Bark Beetle showing significantly improved agreement with independent manual reference data compared to IDS* ***(see Fig. 9)****.*
>
> ***Improving disturbance mapping allows for constraining downstream calculations, such as quantifying actual forest loss, estimating carbon emissions by disturbance type, and correcting carbon loss uncertainty (Harris et al., 2016). The new hybrid dataset proposed here also addresses key challenges in forest disturbance classification and prediction algorithms, particularly those based on machine- and deep learning models (ML and DL, respectively) (Rodríguez Paulino et al., 2024). These data-intensive methods require large amounts of reliable, agent-specific training data. By reducing spatial and temporal noise in IDS, S1DM***

provides a publicly available, standardized benchmark dataset for training and validating methods for forest disturbance classification and prediction. Combined with complementary remote sensing products such as Sentinel-2 or Landsat, this approach has the potential to improve disturbance detection and classification, including the differentiation of subtle disturbance types, such as between different insect disturbance agents or between wind and bark beetle compound events. Previous studies have demonstrated that such distinctions are possible at local or site scales for individual events using optical or multi-sensor data (Eklundh et al., 2009; Meddens et al., 2012; Senf et al., 2015; Negrón-Juárez et al., 2018; Senf and Seidl, 2018; Candotti et al., 2022). However, the systematic and large-scale application of these approaches remains challenging due to limited reference data and the complexity of disturbance dynamics. In this context, S1DM provides a resource that supports the development of scalable, automated disturbance classification frameworks and facilitates the study of ecologically relevant disturbance processes across broader spatial and temporal extents.

Despite these advances, some uncertainties remain, particularly regarding the precise timing and duration of gradual disturbances. Future work could explore ***integrating*** *additional data streams, such as Sentinel-2 time series, and advanced learning strategies to further reconcile information from IDS, S1DM, and optical observations, particularly* ***in*** *cases where neither dataset fully captures the disturbance signal.*

Overall, our approach offers a scalable and substantially less labor-intensive alternative to manual labeling and is readily transferable to other regions and legacy inventories ***(FORWIND; Forzieri et al., 2020; DEFID2; Forzieri et al. 2023).*** *By improving the reliability of disturbance reference data, this work lays the groundwork for more robust data-driven analyses of forest disturbance dynamics across large spatial and temporal scales."*

[General formatting:] Please ensure consistent paragraph indentation (either all first-line indents or none) and uniform spacing between paragraphs throughout the manuscript. Currently, indentation is inconsistent, and spacing varies. Additionally, fix inconsistencies such as missing figure citations in the text (e.g., L427 "Panel a" without "Figure 5"), incorrect numbering in subsections (e.g., "5.1.2 (2)" in the Discussion), and other formatting artifacts. A thorough clean-up is essential for readability.

R2C10: Please ensure consistent paragraph indentation (either all first-line indents or none) and uniform spacing between paragraphs throughout the manuscript. Currently, indentation is inconsistent, and spacing varies. Additionally, fix inconsistencies such as missing figure citations in the text (e.g., L427 "Panel a" without "Figure 5"), incorrect numbering in subsections (e.g., "5.1.2 (2)" in the Discussion), and other formatting artifacts. A thorough clean-up is essential for readability.

We thank the reviewer for pointing out these formatting inconsistencies. We have thoroughly revised the entire manuscript, correcting paragraph indentation, spacing, figure citations, subsection numbering, and other formatting issues. The manuscript should now be consistent and fully aligned with standard formatting conventions.

Minor comments:

*L16: The acronym “Sentinel-1 Disturbance Mapping (S1DM)” should be introduced when it first appears (around line 12), rather than later in the abstract. Define the term before using the abbreviation.

Thank you for noticing. We have now introduced the acronym appropriately at first use in Line 11.

> L. 11: *“We present a novel approach for refining disturbance classification labels by combining IDS with Sentinel-1 radar backscatter change detection to produce a new reference dataset, Sentinel-1 Disturbance Mapping (S1DM). ”*

*L40: Clarify early in the paragraph that these percentages refer to the coverage of disturbance reports (i.e., data availability) rather than the actual affected forest area, to avoid possible misinterpretation.

We have revised sentences in L39–41 to clarify this point and it now reads:

> *L. 40-44: “However, the quality and extent of reporting vary **significantly** across regions. **For example, FAO (2020), shows that the availability of disturbance reports (i.e., data availability) differs considerably by disturbance type and geographic region. Between 2002 to 2016,** data availability for insect disturbance reports covered 98% of forested areas in North and Central America and 86% in Europe, but only 45% in Asia (FAO, 2020).”*

*L58: The phrase “more recently, the Sentinel fleet” should specify the approximate launch period to better clarify the temporal contrast with MODIS and Landsat.

Thank you for pointing this out. We agree with the reviewer that the original sentence was not sufficiently clear and assumed too much prior knowledge. To improve the clarity of the manuscript, we have rewritten the sentence, which now reads:

> *L. 58-63: “Remote sensing provides spatially and temporally consistent and cost-effective solutions for tracking forest conditions across large areas and extended periods. Several long-running satellite missions and sensors support these efforts, including MODIS (operational since 1999; Justice et al., 1998), the Landsat program (initiated with Landsat 1 in 1972 and most recently continued with Landsat 8 launched in 2013; Markham and Helder, 2012; Hansen and Loveland, 2012), and the Sentinel satellite fleet operated by ESA, with multiple satellite missions launched between 2014 and 2025 (ESA, 2015).”*

*The study area description in Section 2.1 lacks quantitative climate data (e.g., precipitation, temperature) essential for contextualizing disturbances. Please add a brief summary.

Thank you for this suggestion. We agree that quantitative climate information is important for contextualizing forest disturbances. We have therefore added a brief summary of temperature and precipitation patterns in Section 2.1 to better characterize the climatic conditions of the study area.

> L. 129-137: "*The study region spans a broad climatic gradient typical of the southeastern United States. Mean annual temperatures generally range from approximately 12–14 °C in the northern and higher-elevation areas to 20–22 °C in the southern coastal and lowland regions. Annual precipitation is relatively high across the region, typically ranging from about 1,100 to 1,600 mm, with higher rainfall in coastal zones and the Appalachian Highlands. The region is characterized by humid subtropical conditions in the south and east, transitioning to more temperate climates at higher elevations and latitudes. In addition, large portions of the study area lie within the North Atlantic hurricane track, making the region frequently exposed to tropical storms and hurricanes that bring strong winds and heavy rainfall, particularly in coastal and lowland areas. These climatic gradients and extreme weather events strongly influence forest composition, productivity, and disturbance regimes, including insect outbreaks, storm damage, and drought-related stress.*"

*L153: The Tree Canopy Cover (TCC) dataset is mentioned for the first time without introduction. Consider briefly referencing it in the Introduction when discussing validation approaches.

We thank the reviewer for this suggestion. We have now briefly introduced the Tree Canopy Cover (TCC) dataset in the Introduction when discussing validation approaches. We have added the following sentences (highlighted in bold) to the paragraph:

> L. 109-111: "***To support validation and quality control, we additionally incorporate a Tree Canopy Cover (TCC) dataset, which provides independent high-resolution estimates of forest canopy cover, and use high-resolution Planet imagery with manual labeling to assess whether this approach leads to significantly improved disturbance detection.***"

*L157: This sentence repeats IDS basics already covered in the Introduction.

We appreciate the reviewer's comment. While the basics of the IDS dataset are indeed introduced in the Introduction, we believe it is important to provide a full description of the dataset in the Data section as well, where we formally introduce all datasets used in this study. This ensures that readers have all the necessary details in the context of our analysis, without needing to refer back to the Introduction.

*L216: The text states that TCC data are available for 2015–2020, but Table 1 lists only 2017. Please clarify whether only 2017 data were used (and why), whether it is a time series, or describe the selection criteria to avoid confusion.

Thank you very much for spotting this. This was a carryover from previous drafts, and we have now updated Table 1 to correctly list the time range as 2015–2020.

*L235: PlanetScope is introduced here without prior mention in the Introduction or Table 1. Please clarify its specific role in the analysis (e.g., manual validation, comparison with S1DM/IDS) and explain why it was not included in Table 1 alongside the key datasets.

We thank the Reviewer for highlighting this inconsistency. We agree that it was an oversight not to list Planet alongside the other datasets in Table 1. We have therefore updated Table 1 to include Planet, adding a dedicated row for each dataset specifying its purpose in our study. Additionally, we have added a sentence to L.235 to clarify PlanetScope's role, consistent with how we describe the other datasets in the previous subsection. The sentence now reads:

> *L. 232-233: "**We used Planet data in this study to manually validate the new dataset and statistically assess whether the methodology produced a significantly improved product compared to IDS.**"*

We also state the inclusion of Planet satellite data in our introduction:

> L. 109-111: "***To support validation and quality control, we additionally incorporate a Tree Canopy Cover (TCC) dataset, which provides independent high-resolution estimates of forest canopy cover, and use high-resolution Planet imagery with manual labeling to assess whether this approach leads to significantly improved disturbance detection.***"

The updated table now shows:

Table 1: Key characteristics of the datasets used in this study: the Insect and Disease Survey (IDS), Sentinel-1 Change Detection (S1CD), Tree Canopy Cover (TCC), **and Planet imagery**. The table summarizes the information content, data format, spatial resolution, temporal availability, **and role of each dataset in the analysis**.

	IDS	S1CD	TCC	**Planet**
Information	Forest damage and mortality caused by various disturbance agents	Information on structural change on the Earth's surface based on Sentinel-1	Tree canopy cover for CONUS in 2017	**High-resolution optical imagery capturing surface reflectance and**

				vegetation dynamics
Data	Spatial data structured in points and polygons[a] with associated attributes documenting forest health and disturbances.	Raster[b] data in binary format (1s and 0s) indicating structural change status.	Raster[b] data in GeoTIFF format representing tree canopy cover as a percentage.	**Raster[b] imagery in multiple spectral bands (e.g., RGB, NIR) at high spatial resolution**
Spatial resolution	0.5 m² – 62,231 km² polygons[a]	20 x 20 m raster	30 x 30 m raster	**~3–5 m (depending on sensor)**
Temporal availability	Yearly data since 1997	Yearly data from 2016 to 2021	**2016-2020**	**Daily to weekly, depending on location and cloud cover**
Purpose in Study	**Provides agent-specific disturbance information for comparison and analysis**	**Detects structural changes in forest canopy to cross-validate IDS and quantify disturbances**	**Identifies forested areas and canopy cover to exclude non-forest disturbances**	**Serves as an independent optical validation source**

[a] Polygons are vector-based, representing features with precise boundaries using points, lines, and polygons.

[b] Raster data is pixel-based, representing space as a grid of cells with values. A disturbance (single pixel or multiple pixels) is referred to as a *patch*.

*L411: Standardize result reporting (e.g., all as "% IDS events with S1DM match") for consistency across disturbance types. Also clarify in the text whether percentages refer to totals across all years (matching absolute numbers in the caption: wind 478 IDS/469 S1DM, etc.).

We thank the reviewer for this helpful comment. Indeed, the text previously alternated between reporting percentages of corresponding and non-corresponding events, which could be confusing for the reader. We have updated the section, including Figure 2.1 (Manuscript Figure 4) requested in **RC2**, which highlights omission percentages more clearly.

Please refer to **RC2** for the changes.

*L427: The text refers to “Panel a” without specifying “Figure 5.” Please include the full figure number (e.g., “Figure 5, Panel a”) and correct similar minor inconsistencies throughout.

Thank you for pointing this out. Upon reviewing the manuscript, we found several instances in which figure references were inconsistently formatted. We have corrected the specific error in L427 and revised the rest of the manuscript to follow a consistent notation throughout, as you suggested (e.g., Fig. 6a).

*L428: The figure interrupts the paragraph. Please follow standard formatting by placing figures at the beginning or end of relevant paragraphs or sections.

Thank you for pointing out the formatting issue. We have adjusted the figure placement to conform to standard formatting conventions, now positioning it at the end of the relevant paragraph.

*Figure 8: The legend shows a black star, but the plot uses an asterisk (*). Please standardize the symbols.

Thank you for noting the inconsistency between the legend and the plot. We have updated Figure 2.2 (Fig. 9 in manuscript) below, so that the legend and plot now use the same symbols, ensuring full consistency.

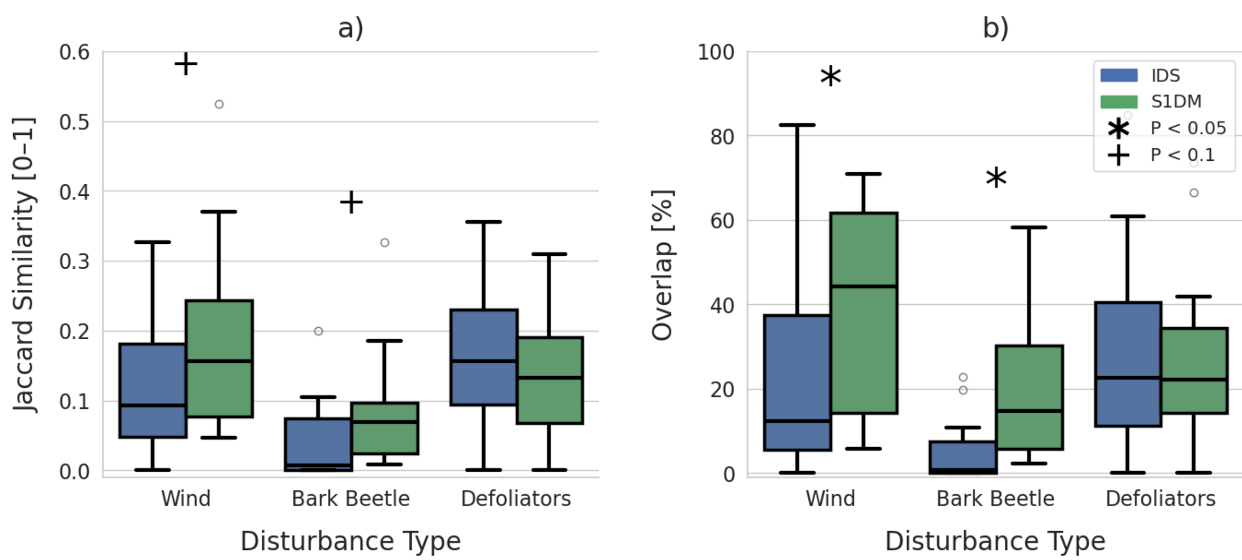

*Figure 2.2: Comparison of agreement between IDS and S1DM disturbance maps with manually delineated reference polygons. Panel (a) shows the Jaccard similarity, and panel (b) shows area overlap (in percentage), both displayed as box plots. Statistical significance from **a one-sided Wilcoxon rank test** is indicated by a black star for p < 0.05 and a black cross for p < 0.1.*

*Sections “5.1.1 (1) Location and delineation of disturbance patches” and “5.1.2 (2) Timing of disturbances:” Please correct subsection numbering errors (remove numbers in parentheses).

We thank the Reviewer for pointing this out. We have removed the numbers in parentheses and corrected the subsection numbering in Sections 5.1.1 and 5.1.2.

References:

[1] Andresini, G., Appice, A., & Malerba, D. (2024). A deep semantic segmentation approach to map forest tree dieback in Sentinel-2 data. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 17, 17075–17086. https://doi.org/10.1109/JSTARS.2024.3460981

[2] Andrus, R. A., Hicke, J. A., & Meddens, A. J. H. (2025). Spatiotemporal characteristics of tree mortality from bark beetle outbreaks vary within and among bark beetle-host tree associations in the western United States. Forest Ecology and Management, 576, 122382. https://doi.org/10.1016/j.foreco.2024.122382

[3] Baumann, M., Ozdogan, M., Wolter, P. T., Krylov, A., Vladimirova, N., & Radeloff, V. C. (2014). Landsat remote sensing of forest windfall disturbance. Remote Sensing of Environment, 143, 171–179. https://doi.org/10.1016/j.rse.2013.12.020

[4] Bárta, V., Lukeš, P., & Homolová, L. (2021). Early detection of bark beetle infestation in Norway spruce forests of Central Europe using Sentinel-2. International Journal of Applied Earth Observation and Geoinformation, 100, 102335. https://doi.org/10.1016/j.jag.2021.102335

[5] Borlaf-Mena, I., Santoro, M., Villard, L., Badea, O., & Tanase, M. (2020). Investigating the impact of digital elevation models on Sentinel-1 backscatter and coherence observations. Remote Sensing, 12(18), 3016. https://doi.org/10.3390/rs12183016

[6] Bruggisser, M., Dorigo, W., Dostálová, A., Hollaus, M., Navacchi, C., Schlaffer, S., & Pfeifer, N. (2021). Potential of Sentinel-1 C-band time series to derive structural parameters of temperate deciduous forests. Remote Sensing, 13(4), 798. https://doi.org/10.3390/rs13040798

[7] Canadell, J. G., Meyer, C. P. (M.), Cook, G. D., Dowdy, A., Briggs, P. R., Knauer, J., Pepler, A., & Haverd, V. (2021). Multi-decadal increase of forest burned area in Australia is linked to climate change. Nature Communications, 12(1), 6921. https://doi.org/10.1038/s41467-021-27225-4

[8] Candotti, A., De Giglio, M., Dubbini, M., & Tomelleri, E. (2022). A Sentinel-2 based multi-temporal monitoring framework for wind and bark beetle detection and damage mapping. Remote Sensing, 14(23), 6105. https://doi.org/10.3390/rs14236105

[9] Chen, X., Zhao, W., Chen, J., Qu, Y., Wu, D., & Chen, X. (2021). Mapping large-scale forest disturbance types with multi-temporal CNN framework. Remote Sensing, 13(24), 5177. https://doi.org/10.3390/rs13245177

[10] Cohen, W. B., Yang, Z., & Kennedy, R. E. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 2. Time series analysis. Remote Sensing of Environment, 114, 2897–2910. https://doi.org/10.1016/j.rse.2010.07.010

[11] Coleman, T. W., Graves, A. D., Heath, Z., Flowers, R. W., Hanavan, R. P., Cluck, D. R., & Ryerson, D. (2018). Accuracy of aerial detection surveys for mapping insect and disease disturbances in the United States. Forest Ecology and Management, 430, 321–336. https://doi.org/10.1016/j.foreco.2018.08.020

[12] Cremer, F., Urbazaev, M., Cortes, J., Truckenbrodt, J., Schmullius, C., & Thiel, C. (2020). Potential of recurrence metrics from Sentinel-1 time series for deforestation mapping. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13, 5233–5240. https://doi.org/10.1109/JSTARS.2020.3019333

[13] Cremer, F., Gans, F., Cortes, J., & Thiel, C. (2023). Mapping forest loss in Europe with Sentinel-1. In P. Soille, S. Lumnitz, & S. Albani (Eds.), Proceedings of the 2023 Conference on Big Data from Space (BiDS'23): From foresight to impact (pp. 361–364). Publications Office of the European Union. https://doi.org/10.2760/46796

[14] Eifler, L., Müller, F., & Bastos, A. (2026). Evaluating the consistency of forest disturbance datasets in continental USA. Biogeosciences, 23(4), 1291–1325. https://doi.org/10.5194/bg-23-1291-2026

[15] Eklundh, L., Johansson, T., & Solberg, S. (2009). Mapping insect defoliation in Scots pine with MODIS time-series data. Remote Sensing of Environment, 113(7), 1566–1573. https://doi.org/10.1016/j.rse.2009.03.008

[16] European Commission. (2023). Proposal for a regulation of the European Parliament and of the Council on a monitoring framework for resilient European forests (COM(2023) 728 final). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52023PC0728

[17] European Space Agency (ESA). (2015). ESA's Living Planet Programme: Scientific achievements and future challenges – Scientific context of the Earth observation science strategy for ESA (ESA SP-1329/2). https://esamultimedia.esa.int/multimedia/publications/SP-1329_2/SP-1329-2.pdf

[18] Food and Agriculture Organization of the United Nations. (2020). Global forest resources assessment 2020 (FRA2020). https://doi.org/10.4060/ca9825en

[19] Forest Health Detection Survey: GIS HandbookandData Conformity StandardsVersion 1.0, Forest Service U.S. Department of Agriculture, January 2022

[20] Forzieri, G., Pecchi, M., Girardello, M., Mauri, A., Klaus, M., Nikolov, C., Rüetschi, M., Gardiner, B., Tomaštík, J., Small, D., Nistor, C., Jonikavicius, D., Spinoni, J., Feyen, L., Giannetti, F., Comino, R., Wolynski, A., Pirotti, F., Maistrelli, F., Savulescu, I., Wurpillot-Lucas, S., Karlsson, S., Zieba-Kulawik, K., Strejczek-Jazwinska, P., Mokroš, M., Franz, S., Krejci, L., Haidu, I., Nilsson, M., Wezyk, P., Catani, F., Chen, Y.-Y., Luyssaert, S., Chirici, G., Cescatti, A., & Beck, P. S. A. (2020). A spatially explicit database of wind disturbances in European forests

over the period 2000–2018. Earth System Science Data, 12(1), 257–276. https://doi.org/10.5194/essd-12-257-2020

[21] Forzieri, G., Dutrieux, L. P., Elia, A., Eckhardt, B., Caudullo, G., Álvarez Taboada, F., Andriolo, A., Bălăcenoiu, F., Bastos, A., Buzatu, A., Castedo Dorado, F., Dobrovolný, L., Duduman, M.-L., Fernandez-Carrillo, A., Cescatti, A., Beck, P. S. A., & Forzieri, G. (2023). The database of European forest insect and disease disturbances: DEFID2. Global Change Biology, 00, 1–30. https://doi.org/10.1111/gcb.16912

[22] Gibson, R., Danaher, T., Hehir, W., & Collins, L. (2020). A remote sensing approach to mapping fire severity in south-eastern Australia using Sentinel-2 and random forest. Remote Sensing of Environment, 240, 111702. https://doi.org/10.1016/j.rse.2020.111702

[23] Hall, R. J., Castilla, G., White, J. C., Cooke, B. J., & Skakun, R. S. (2016). Remote sensing of forest pest damage: A review and lessons learned from a Canadian perspective. The Canadian Entomologist, 148(S1), S296–S356. https://doi.org/10.4039/tce.2016.11

[24] Hansen, M. C., Potapov, P. V., Moore, R., Hancher, M., Turubanova, S. A., Tyukavina, A., Thau, D., Stehman, S. V., Goetz, S. J., Loveland, T. R., Kommareddy, A., Egorov, A., Chini, L., Justice, C. O., & Townshend, J. R. G. (2024). Global Forest Change 2000–2024 data download. Google Earth Engine. https://storage.googleapis.com/earthenginepartners-hansen/GFC-2024-v1.12/download.html

[25] Hansen, M. C., & Loveland, T. R. (2012). A review of large area monitoring of land cover change using Landsat data. Remote Sensing of Environment, 122, 66–74. https://doi.org/10.1016/j.rse.2011.08.024

[26] Harris, N. L., Hagen, S. C., Saatchi, S. S., Pearson, T. R. H., Woodall, C. W., Domke, G. M., Braswell, B. H., Walters, B. F., Brown, S., Salas, W., Fore, A., & Yu, Y. (2016). Attribution of net carbon change by disturbance type across forest lands of the conterminous United States. Carbon Balance and Management, 11(1), 24. https://doi.org/10.1186/s13021-016-0066-5

[27] Hawryło, P., Bednarz, B., Wężyk, P., & Szostak, M. (2018). Estimating defoliation of Scots pine stands using machine learning methods and vegetation indices of Sentinel-2. European Journal of Remote Sensing, 51(1), 194–204. https://doi.org/10.1080/22797254.2017.1417745

[28] Heaton, R., Song, B., Williams, T., Conner, W., Baucom, Z., & Williams, B. (2023). Twenty-seven year response of South Carolina coastal plain forests affected by Hurricane Hugo. Plants, 12(4), 691. https://doi.org/10.3390/plants12040691

[29] Helmer, E. H., Ruzycki, T. S., Wunderle Jr., J. M., Vogesser, S., Ruefenacht, B., Kwit, C., Brandeis, T. J., & Ewert, D. N. (2010). Mapping tropical dry forest height, foliage height profiles and disturbance type and age with a time series of cloud-cleared Landsat and ALI image

mosaics to characterize avian habitat. Remote Sensing of Environment, 114, 2457–2473. https://doi.org/10.1016/j.rse.2010.05.021

[30] Hicke, J. A., Xu, B., Meddens, A. J. H., & Egan, J. M. (2020). Characterizing recent bark beetle-caused tree mortality in the western United States from aerial surveys. Forest Ecology and Management, 475, 118402. https://doi.org/10.1016/j.foreco.2020.118402

[31] Hirschmugl, M., Gallaun, H., Dees, M., Datta, P., Deutscher, J., Koutsias, N., & Schardt, M. (2017). Methods for mapping forest disturbance and degradation from optical Earth observation data: A review. Current Forestry Reports, 3(1), 32–45. https://doi.org/10.1007/s40725-017-0047-2

[32] Jaccard, P. (1902). Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bulletin de la Société Vaudoise des Sciences Naturelles, 37, 547–579.

[33] Justice, C. O., Vermote, E., Townshend, J. R. G., Defries, R., Roy, D. P., Hall, D. K., Salomonson, V. V., Privette, J. L., Riggs, G., Strahler, A., Lucht, W., Myneni, R. B., Knyazikhin, Y., Running, S. W., Nemani, R. R., Wan, Z., Huete, A. R., van Leeuwen, W., Wolfe, R. E., Giglio, L., Muller, J., Lewis, P., & Barnsley, M. J. (1998). The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research. IEEE Transactions on Geoscience and Remote Sensing, 36(4), 1228–1249. https://doi.org/10.1109/36.701075

[34] Kautz, M., Meddens, A. J. H., Hall, R. J., & Arneth, A. (2017). Biotic disturbances in Northern Hemisphere forests: A synthesis of recent data, uncertainties and implications for forest monitoring and modelling. Global Ecology and Biogeography, 26, 533–552. https://doi.org/10.1111/geb.12558

[35] Kellndorfer, J. (2019). Using SAR data for mapping deforestation and forest degradation. In A. I. Flores-Anderson, K. E. Herndon, R. B. Thapa, & E. Cherrington (Eds.), Synthetic aperture radar (SAR) handbook: Comprehensive methodologies for forest monitoring and biomass estimation (pp. 47–66). SERVIR Global Science Coordination Office, National Space Science and Technology Center.

[36] Kennedy, R. E., Yang, Z., & Cohen, W. B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr — Temporal segmentation algorithms. Remote Sensing of Environment, 114(12), 2897–2910. https://doi.org/10.1016/j.rse.2010.07.008

[37] Kislov, D. E., Korznikov, K. A., Altman, J., Vozmishcheva, A. S., & Krestov, P. V. (2021). Extending deep learning approaches for forest disturbance segmentation on very high-resolution satellite images. Remote Sensing in Ecology and Conservation, 7(3), 355–368. https://doi.org/10.1002/rse2.194

[38] Markham, B. L., & Helder, D. L. (2012). Forty-year calibrated record of earth-reflected radiance from Landsat: A review. Remote Sensing of Environment, 122, 30–40. https://doi.org/10.1016/j.rse.2011.06.026

[39] Marwan, N., Romano, M. C., Thiel, M., & Kurths, J. (2007). Recurrence plots for the analysis of complex systems. Physics Reports, 438(5–6), 237–329. https://doi.org/10.1016/j.physrep.2006.11.001

[40] McDowell, N. G., Coops, N. C., Beck, P. S. A., Chambers, J. Q., Gangodagamage, C., Hicke, J. A., Huang, C.-y., Kennedy, R., Krofcheck, D. J., Litvak, M., Meddens, A. J. H., Muss, J., Negrón-Juárez, R., Peng, C., Schwantes, A. M., Swenson, J. J., Vernon, L. J., Williams, A. P., Xu, C., Zhao, M., & Running, S. W. (2015). Global satellite monitoring of climate-induced vegetation disturbances. Trends in Plant Science, 20(2), 114–123. https://doi.org/10.1016/j.tplants.2014.12.001

[41] Meddens, A. J. H., Hicke, J. A., & Ferguson, C. A. (2012). Spatiotemporal patterns of observed bark beetle-caused tree mortality in British Columbia and the western United States. Ecological Applications, 22, 1876–1891.

[42] Migliavacca, M., Grassi, G., Bastos, A., Ceccherini, G., Ciais, P., Janssens-Maenhout, G., Lugato, E., Mahecha, M., Novick, K., Peñuelas, J., Pilli, R., Reichstein, M., Avitabile, V., Beck, P. S. A., Barredo, J. I., Forzieri, G., Herold, M., Korosuo, A., Mansuy, N., Mubareka, S., Orth, R., Rougieux, P., & Cescatti, A. (2025). Securing the forest carbon sink for the European Union's climate ambition: Scientific action to anticipate the evolution of the forest carbon sink. Nature, 643(8074), 1203–1213. https://doi.org/10.1038/s41586-025-08967-3

[43] Negrón-Juárez, R. I., Holm, J. A., Marra, D. M., Rifai, S. W., Riley, W. J., Chambers, J. Q., Koven, C. D., Knox, R. G., McGroddy, M. E., Vittorio, A. V. Di, Urquiza-Muñoz, J., Tello-Espinoza, R., Muñoz, W. A., Ribeiro, G. H. P. M., & Higuchi, N. (2018). Vulnerability of Amazon forests to storm-driven tree mortality. Environmental Research Letters, 13(5), 054021. https://doi.org/10.1088/1748-9326/aabe9f

[44] Oeser, J., Pflugmacher, D., Senf, C., Heurich, M., & Hostert, P. (2017). Using intra-annual Landsat time series for attributing forest disturbance agents in Central Europe. Forests, 8(7), 251. https://doi.org/10.3390/f8070251

[45] Planet Labs PBC. (2025). *PlanetScope satellite imagery* [dataset]. Accessed 2025, from https://www.planet.com/products/planetscope/

[46] Rodríguez Paulino, E., Schlerf, M., Röder, A., Stoffels, J., & Udelhoven, T. (2024). Forest disturbance characterization in the era of earth observation big data: A mapping review. International Journal of Applied Earth Observation and Geoinformation, 128, 103755. https://doi.org/10.1016/j.jag.2024.103755

[47] Schleeweis, K. G., Moisen, G. G., Schroeder, T. A., Toney, C., Freeman, E. A., Goward, S. N., Huang, C., & Dungan, J. L. (2020). US national maps attributing forest change: 1986–2010. Forests, 11(6), 653. https://doi.org/10.3390/f11060653

[48] Schroeder, T. A., Schleeweis, K. G., Moisen, G. G., Toney, C., Cohen, W. B., Freeman, E. A., Yang, Z., & Huang, C. (2017). Testing a Landsat-based approach for mapping disturbance causality in U.S. forests. Remote Sensing of Environment, 195, 230–243. https://doi.org/10.1016/j.rse.2017.03.033

[49] Senf, C., Pflugmacher, D., Wulder, M. A., & Hostert, P. (2015). Characterizing spectral–temporal patterns of defoliator and bark beetle disturbances using Landsat time series. Remote Sensing of Environment, 170, 166–177. https://doi.org/10.1016/j.rse.2015.09.019

[50] Senf, C., & Seidl, R. (2018). Natural disturbances are spatially diverse but temporally synchronized across temperate forest landscapes in Europe. Global Change Biology, 24(3), 1201–1211. https://doi.org/10.1111/gcb.13897

[51] Senf, C., & Seidl, R. (2021). Storm and fire disturbances in Europe: Distribution and trends. Global Change Biology, 27(15), 3605–3619. https://doi.org/10.1111/gcb.15679

[52] Shi, C., Zuo, X., Zhang, J., Zhu, D., Li, Y., & Bu, J. (2024). Accuracy assessment of geometric-distortion identification methods for Sentinel-1 synthetic aperture radar imagery in highland mountainous regions. Sensors, 24(9), 2834. https://doi.org/10.3390/s24092834

[53] Stahl, A. T., Andrus, R., Hicke, J. A., Hudak, A. T., Bright, B. C., & Meddens, A. J. H. (2023). Automated attribution of forest disturbance types from remote sensing data: A synthesis. Remote Sensing of Environment, 285, 113416. https://doi.org/10.1016/j.rse.2022.113416

[54] United States Department of Agriculture, Forest Service. (2015). *Major Forest Insect and Disease Conditions in the United States: 2015*. Washington, DC: U.S. Department of Agriculture, Forest Service.

[55] Urquiza-Muñoz, J. D., Trumbore, S., Negrón-Juárez, R. I., Feng, Y., Brenning, A., Vasquez-Parana, C. M., & Marra, D. M. (2024). Increased occurrence of large-scale windthrows across the Amazon Basin. AGU Advances, 5(6), e2023AV001030. https://doi.org/10.1029/2023AV001030

[56] USDA Forest Service (2018). *Aerial Detection Survey Results, 2017*. Forest Health Protection, U.S. Department of Agriculture.

[57] USDA Forest Service: Aerial Survey Geographic Information System Handbook: Sketchmaps to Digital Geographic Information, USDAForest Service, Forest Health Monitoring Program, State and Private Forestry, Forest Health Protection, Aerial Survey GIS Handbook, 2005.

[58] USDA Forest Service: Forest Health Detection Survey GIS Handbook and Data Conformity Standards, USDA Forest Service, Forest Health Monitoring Program, 2022.

[59] Vanguri, R., Laneve, G., & Hościło, A. (2024). *Mapping forest tree species and its biodiversity using EnMAP hyperspectral data along with Sentinel-2 temporal data: An approach of tree species classification and diversity indices*. **Ecological Indicators**, 153, 112671. https://doi.org/10.1016/j.ecolind.2024.112671

[60] Verbesselt, J., Hyndman, R. J., Newnham, G., & Culvenor, D. (2010). Detecting trend and seasonal changes in satellite image time series. Remote Sensing of Environment, 114, 106–115. https://doi.org/10.1016/j.rse.2009.08.014

[61] Viana-Soto, A., & Senf, C. (2025). The European Forest Disturbance Atlas: A forest disturbance monitoring system using the Landsat archive. Earth System Science Data, 17(6), 2373–2404. https://doi.org/10.5194/essd-17-2373-2025

[62] Woudenberg, S. W., Conkling, B. L., O'Connell, B. M., LaPoint, E. B., Turner, J. A., & Waddell, K. L. (2010). The Forest Inventory and Analysis Database: Database description and users manual version 4.0 for Phase 2. U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. https://doi.org/10.2737/RMRS-GTR-245

[63] Zhu, Z., & Woodcock, C. E. (2014). Continuous change detection and classification of land cover using all available Landsat data. Remote Sensing of Environment, 144, 152–171. https://doi.org/10.1016/j.rse.2014.01.011