Evaluating the consistency of forest disturbance datasets in continental USA
Laura Eifler, Franziska Müller, and Ana Bastos, *EGUsphere*

Response to Reviewer #2

R2C1: I found the manuscript to be generally well written and clear. However, the description of each dataset was unnecessarily long, as detailed methods for each are in the original papers. Removing some of the extraneous details from Section 2 would greatly reduce the size of the text. The methods used were also generally well executed. Using Gaussian distributions to track the lag in disturbance event detection between datasets was a clever method that is simple, intuitive, and effective.
However, I also have a few major issues and suggestions for the authors:
We thank the reviewer for the evaluation and constructive comments. Below, we provide a point-by-point reply to the reviewer's comments. In line with the suggestion, the data and method sections are reduced now, we have shortened the text and only include the information about each dataset that is directly relevant for this study.

Paper Framing/Motivation
R2C2: First, I have concerns about the framing. As a reader, I found myself wondering what the "so what?" of the paper was. The current paper framing suggests that there is a need to compare ground vs remotely sensed datasets for consistency and accuracy. However, the selection of datasets for comparison seems arbitrary. The ground datasets are part of a regional network of concentrated sites with a goal of intensively sampling the entire continental United States, while the remote sensing datasets included are at a global scale. To me, this creates a **mismatch in scale** that is not rectified within the current framing. Is there a need for better regional assessments of global products? If this is not the motivation, then I think a better comparison would be in using remote sensing datasets that have been calibrated or developed regionally, such as the **North American Forest Dynamics (NAFD) Forest Loss Attribution dataset (Schleeweis et al. 2020)**, or the Monitoring Trends in Burn Severity (MTBS) product. So I am left wondering what this analysis is hoping to accomplish. I could see a reasonable argument for the following:

1. Providing a regional assessment of global remote sensing products using field inventories as a benchmark. – with hansen
2. Evaluating dataset performance or appropriate use in different contexts. For instance, determining the scenarios in which using field vs remote sensing datasets (or a combination) might be more useful.

In both cases, it seems the paper needs reconfiguring.
Thank you for your thoughtful feedback. We understand your concern regarding the mismatch in scale between the ground-based and remote sensing datasets, as well as the framing of the paper. You raise an important point about the need for better regional assessments of global products, and we agree that this is a significant consideration.
Our main motivation aligns in between both points proposed by the reviewer. Here, we aim to provide a regional assessment of the uncertainties associated with forest disturbance mapping and agent attribution. Remote-sensing products are an essential part for global studies, but inventory data such as FIA and IDS provide unique information that can be used as a benchmark for remote-sensing datasets. However, as we show here, even these data can show important disagreements between themselves, making the selection of a unique benchmark difficult. Such uncertainties are inherent to the fact that there are different collection methods and revisit times, different definitions and groupings of agent types, and other factors, such as property status, that we consider in our study. By comparing the spatial and temporal agreement between different

datasets (inventory data with remote-sensing and inventory between themselves) we aim to quantify the uncertainty associated with different approaches and, when possible, the sources of such uncertainties (e.g. Figures 3 - 7).

In light of the reviewer's feedback, we will reframe the paper to better clarify our motivation and discuss how regional assessments of global products can improve disturbance monitoring and inform future decision-making. Additionally, we will incorporate the North American Forest Dynamics (NAFD) Forest Loss Attribution dataset (Schleeweis et al., 2020) into the analysis, as suggested, to strengthen the comparison with other remotely sensed products developed at a regional scale.

Major issues with Methods

R2C3: In addition to a reframing and proper motivation, A clear explanation for the decision to compare these 4 datasets in particular is needed.

We agree with the reviewer that the motivation of the study can be improved. The text has been revised according to our answer to R2C2 and we have added a paragraph in the Introduction explaining the decision using the four (five in the revised manuscript) datasets:

> *"Here, we aim to compare the consistency between state-of-the-art and openly available datasets on forest disturbances from two broad groups: ground survey and inventory data (referred here as "inventory"), and remote-sensing datasets. The continental USA is one of the global regions with a richer collection of such datasets covering a long period and overlapping in time: the FIA, ITMN and IDS (inventory), and GFC and NAFD (remote-sensing). Inventories provide rich, detailed information on disturbance location, extent, timing, and agents. However, these datasets are limited in their spatial and temporal coverage and can be affected by various uncertainties arising from sampling strategies, reporting methods, and human error. On the other hand, remote-sensing products like the GFC and NAFD offer large-scale, spatially and temporally consistent maps of forest disturbances. While these datasets excel in covering vast areas, their accuracy is influenced by factors such as the spatial resolution of the imagery, the intensity of disturbances, and the underlying forest structure.*
>
> *The complementary nature of these datasets is the key factor in their selection for this study. Inventory data, with their detailed ground-level observations, are essential for accurately attributing disturbances to specific agents and for validating remote-sensing products.*
>
> *Our goal is to quantify the robustness of the information on disturbance year and respective agents, and to identify advantages and potential shortcomings and inaccuracies of these datasets. By evaluating widely used global datasets against more detailed ground-based inventories, we aim to assess the reliability of large-scale forest disturbance monitoring and identify potential biases or inconsistencies. This approach serves as a case study, offering insights into the general challenges of validating global datasets with inventory data."*

R2C4: If disturbance agent is a foci of the paper, I do not understand why IDS is the only remotely sensed disturbance dataset being used as it most heavily represents one disturbance agent. Is insect damage a focus of the paper or are the authors interested in all disturbances? If all, then it would be more appropriate to compare other regional disturbance datasets derived from remote sensing instead of the GFC (see my comment above).

Thanks for the comment. In line with our answer to R2C2, we are interested in evaluating fit-for-purpose of each dataset. IDS is not only remotely-sensed but also relies on field surveys to characterise disturbance agents. To the best of our knowledge it is the most complete spatially-

explicit (i.e. not at plot scale) inventory-based dataset of disturbance agents covering the whole CONUS. In line with R2C2, we will now compare the datasets with NAFD Landsat based dataset. Nevertheless we would like to defend the comparison with GFC in terms of spatial and temporal agreement given its widespread use by the forest monitoring community (e.g., Feng et al., 2022; Tyukavina et al., 2022; Potapov et al., 2020; Li et al., 2017; Hamilton et al., 2016, Antonarakis et al., 2022, Kinnebrew et al., 2022) and the fact that it is continuously updated and publicly available. Therefore, assessing its consistency with inventory data is important to identify weaknesses and strengths and its suitability for specific applications. This includes evaluating how well this open-access dataset reflects forest stand loss and identifying potential disturbance agents responsible for the loss.

R2C5: Why include Alaska in this assessment? FIA collection across Alaska is relatively new and to my knowledge, there has not been an Alaska-wide data release yet. The datasets for Alaska in Figure 1 look to be primarily GFC with very little data from other datasets.

Thank you for raising this point. Initially, we included the data from Alaska as part of the assessment, as it was provided by FIA. However, we agree that the sample size is very small, and given that FIA's data collection in Alaska is still relatively new and there hasn't been a comprehensive data release yet, we will remove the Alaska data from the analysis.

General comments

L65-70 It is difficult to tell which datasets are inventory based vs remote sensing based. Suggest edit and for clarity.

Thank you for pointing this out. The text was edited to clarify the database/ source, as described below:

> *"Specifically, we compare data from two systematic inventories - the Insect and Disease Survey (IDS) data by USDA (Forest Service U.S. Department of Agriculture) and the Forest Inventory and Analysis (FIA) data by USDA (Forest Service U.S. Department of Agriculture) - alongside tree mortality events based on literature review by the International Tree Mortality Network (ITMN, (Hammond et al., 2022)), satellite-based tree cover change data from the Global Forest Watch (GFC, (Hansen et al., 2013)), and the NAFD dataset from Schleeweis et al. (2020) as an additional regional remote sensing product."*

Furthermore we have now added this information to the overview in Table 1.

L73 remotely-sensed rather than remote-sensing

Corrected, thanks.

L84 IDS is defined multiple times. Define once he first time it is mentioned in the main text. Also, a reference and link to the specific dataset documentation or web page would be preferred.

Thank you, that is right. It is now defined in the introduction only and a reference is added in the data section.

L85-90 This section needs rephrasing. Currently large portions of sentences seemed borrowed directly from the web pages describing the dataset and applications.

Thanks, the paragraph is now rephrased:

> *"In the IDS dataset, the continental United States is divided into nine major regions, which are surveyed annually through both aerial detection and ground observations. Data collection relies on applications such as the Digital Mobile Sketch Mapping (DMSM) and the Southern Pine Beetle (SPB) Collector Map.*
>
> *In the DMSM, forest disturbances are documented by sketching the extent of tree injuries and mortality, storing the data as geo-referenced points or polygons. In addition to mapping*

the affected areas, surveyors record key attributes, including tree species, disturbance agents, damage types (e.g., defoliation, mortality, discoloration), and disturbance severity. Geo-referenced base layers, such as aerial photographs, topographic maps, and near-infrared imagery, help track surveyor positions, enhance aerial detection, and prevent duplicate records if damage has been previously documented. The database contains vector data in the form of both polygons and points representing recorded disturbances. This study focuses specifically on polygon data."

L137-138 sentence needs rephrasing for clarity: 78% of global forested area is north of the Equator
We agree with the reviewer and have rephrased this sentence:

"Additionally, there is a strong bias to the Northern Hemisphere, which is explained by the fact that 78% of the global forested area is north of the Equator."

L262. Throughout the writing, the text would benefit from a consistent and clear grouping of the datasets. Are the datasets ground vs remote sensing, point-based vs spatially explicit, etc.? Once a standard language is determined, continuing to group the datasets accordingly across methods, results, and discussion (including figures and tables) would help the reader more easily interpret results.
Thanks for the suggestion. We agree with the reviewer that this would help the comprehensiveness of the text and propose to include a grouping of datasets in point-based vs spatially explicit in the revised manuscript.

L267: Should a word be added here? "Indicating that FIA records disturbance"
Yes, thank you. The sentence was rephrase and now reads:

"The large standard deviation (approximately 7 years for both datasets) suggests that FIA tends to record events earlier than the other datasets, but with a considerable variability."

L353 Fire is a large disturbance component in the United States. It is heavily emphasized in the introduction and discussion, but not reflected in the selected datasets.
We do not fully understand why the reviewer considers that fire as a disturbance agent is heavily emphasized in the introduction and discussion sections, as it is rather mentioned as one of several disturbance agent types and being an integral part of the ecosystem (L18). For example, fire is mentioned seven times, compared to insects (including bark-beetles and defoliators) which are mentioned 14 times.
As in the discussion it serves as an example of a disturbance causing agent which is detected and identified often and reliable, compared to biotic agents such as bark beetles or defoliators which are more uncertain. This is in fact aligned with the purpose of the study, i.e., which datasets are more suitable to study given disturbance types: only few fire records in the IDS and FIA datasets, which can be seen in Figures 4 and 5.

L369-372 Uncertainty in the IDS is well documented in this paper as well: (Cohen etal.,2016)
Thank you for the paper suggestion, we will incorporate it in the discussion.

L410-415 Again, as disturbance detection is emphasized in the study, I do not understand why a global dataset composed of estimated stand replacing disturbances (GFC) was compared with other datasets focused on much more nuanced disturbance patterns.
Thank you for this comment. We understand your concern regarding the inclusion of the GFC dataset, which primarily focuses on stand-replacing disturbances, in comparison with datasets that provide more detailed, nuanced disturbance information. We agree that the granularity of the data differs, but we believe that the comparison helps highlight the strengths and limitations of different

datasets in capturing disturbances of varying sizes and types. This comparison is particularly relevant for identifying gaps or overlaps in disturbance detection capabilities and provides insight into how datasets can be integrated for comprehensive forest disturbance monitoring. In the revised version of our analysis, we will provide additional clarification on how these different datasets complement each other and contribute to a broader understanding of forest disturbance patterns. Also see reply to R2C2.

Figure 1. The IDS polygons in Panel B seem to be plotted over the ITMN and FIA values. I would suggest having the IDS polygons mapped first, followed by the GFC raster data, and the other vector data last.
Thank you for the recommendation, we will improve the figure with the suggested layering.

Figure 4. The figure needs a description of the disturbance agents – e.g. what does BB stand for? Either spell out the disturbance agent on the figure or describe the acronyms in the legend.
Yes, the detailed description of the agents is missing and will be included in the figure caption.

Table 1. FIA coordinates are fuzzed across all land ownership types, not just those on privately owned land. A portion of those on private lands are swapped.
Thank you for pointing that out, it is corrected.

Table 3. Additional description of this table is needed to discern what the rows and columns represent. Also, here and elsewhere, grouping the datasets based on whether they are remote sensing or ground datasets would be helpful (i.e. placing FIA and ITMN next to one another rather than split between IDS and GFC).
Thank you for the suggestion, we corrected the table according to that.

References from comments:
Cohen, W.B., Yang, Z.,Stehman, S.V.,Schroeder, T.A.,Bell, D.M.,Masek, J.G.,et al. (2016). Forest disturbance across the conterminous United States from 1985–2012: The emerging dominance of forest decline. Forest Ecology and Management, 360,242–252. https://doi.org/10.1016/j.foreco.2015.10.042
Schleeweis, K.G., G.G. Moisen, T.A. Schroeder, C. Toney, E.A. Freeman, S.N. Goward, C. Huang, and J.L. Dungan. 2020. US National Maps Attributing Forest Change: 1986–2010. Forests, 11(6), p.653. https://doi.org/10.3390/f11060653

References:
Feng, Y., Zeng, Z., Searchinger, T. D., Ziegler, A. D., Wu, J., Wang, D., ... & Zheng, C. (2022). Doubling of annual forest carbon loss over the tropics during the early twenty-first century. *Nature Sustainability*, *5*(5), 444-451.

Tyukavina, A., Potapov, P., Hansen, M. C., Pickens, A. H., Stehman, S. V., Turubanova, S., Parker, D., Zalles, V., Lima, A., Kommareddy, I., Song, X.-P., Wang, L., & Harris, N. (2022). Global trends of forest loss due to fire from 2001 to 2019. Frontiers in Remote Sensing, 3.

Potapov, P., Hansen, M. C., Kommareddy, I., Kommareddy, A., Turubanova, S., Pickens, A., ... & Ying, Q. (2020). Landsat analysis ready data for global land cover and land cover change mapping. *Remote Sensing*, *12*(3), 426.

Li, Y., Sulla-Menashe, D., Motesharrei, S., Song, X. P., Kalnay, E., Ying, Q., ... & Ma, Z. (2017). Inconsistent estimates of forest cover change in China between 2000 and 2013 from multiple datasets: Differences in parameters, spatial resolution, and definitions. *Scientific reports*, *7*(1), 8748.

Hamilton, S. E., & Casey, D. (2016). Creation of a high spatio-temporal resolution global database of continuous mangrove forest cover for the 21st century (CGMFC-21). *Global Ecology and Biogeography*, *25*(6), 729-738.

Antonarakis, A. S., Pacca, L., & Antoniades, A. (2022). The effect of financial crises on deforestation: A global and regional panel data analysis. *Sustainability Science*, *17*(3), 1037-1057.

Kinnebrew, E., Ochoa-Brito, J. I., French, M., Mills-Novoa, M., Shoffner, E., & Siegel, K. (2022). Biases and limitations of Global Forest Change and author-generated land cover maps in detecting deforestation in the Amazon. *PLoS One*, *17*(7), e0268970.

Hammond, W. M., Williams, A. P., Abatzoglou, J. T., Adams, H. D., Klein, T., López, R., ... & Allen, C. D. (2022). Global field observations of tree die-off reveal hotter-drought fingerprint for Earth's forests. *Nature communications*, *13*(1), 1761.