nature portfolio

Peer Review File

Biophysical factors and management practices are key to shaping forest resilience

Corresponding Author: Dr Zhiyong Liu

This file contains all reviewer reports in order by version, followed by all author rebuttals in order by version.

Attachments originally included by the reviewers as part of their assessment can be found at the end of this file.

Version 0:

Reviewer comments:

Reviewer #1

(Remarks to the Author)
---
title: "Review for manuscript NCOMMS-24-55547"

Overall impression

The manuscript attempts an analysis of the sensitivity of forest resilience to human disturbance and environmental drivers in different forest management categories, arguing that natural forests have higher resilience than managed forests in 60% of the cases. While the subject is of great interest, the manuscript is unfortunately not convincing. The techniques used are inappropriately explained and the limitations of the underlying data are not well exposed. Furthermore, the overall quality and style of the manuscript is not in-line with what is expected in such a journal. References are incomplete, lacking journal names, sometimes pages. There are quite some errors in language that should have been picked up by any spelling corrector. Some phrases seem unfinished (no verb, or no object, see L93). Then technically too, various crucial methodological details are not explained in the main text, making understanding and following the text quite difficult, and undermining the credibility of the claims that the manuscript tries to put forward. Given these various points, the overall recommendation would be rejection. However, if the authors want to improve the manuscript, hopefully this review can help them. Below is a list of more general issues followed a more detailed list of specific points.

General issues

- The abstract gives no clue on the nature of the work. How is resilience quantified? Model simulations? satellite metrics? ground-data?
- Consider that the methodology section comes at the end, so any concept, method, dataset, or abbreviation needs to be introduced before being dropped on the reader without context. A few examples are listed in the detailed comments.
- In the introduction, four warning signals are presented: AR1, variance, λAR1, and λVar. Yet, the results are based only on three, excluding the variance metric. This needs to be motivated or risk feeling like cherry-picking.
- Check that the references to the figures are correct throughout the document. It felt that some were not, which made following along sometimes very confusing. Also, Figures and supplementary Figures should be always presented in order, and all should be referenced in the text.
- Statements based on studies from the literature should probably be moved all to the discussion, as in the results it is not always clear if the statements made are based on own results or the literature. Alternatively, it should be made very clear whenever a reference is cited as basis for a statement, or as support of own conclusions.
- The definition of forest types seems pretty central to the research question, so it should also be central in the text, but it is now only relegated to the supplementary materials. Either tables S1 and/or S2 should be moved to the main manuscript, or the most important forest type definitions should be clearly stated in the text.
- The arguments about the independence between forest management and resilience of the GFMD data should be more central in the discussion, as it is potentially a huge underlying limitation. Therefore, the independent tests on crowdsourcing data should be highlighted more.

Specific points

L17: P/PEF should be defined
L49: correct “bring” to “bringing”.
L60: “However, despite models appearing to predict stably, forest dynamics in observation remain complex and diverse” unclear sentence.
L62: correct “plating” to “planting”.
L64: remove double comma after “Furthermore”; clarify what is meant with “early stable predictions”.
L67: “Recent study report that…” correct singular/plural form (either “A recent study reports that…” or “Recent studies report that…”.
L68: correct “dectect” to “detect”
L69-70: “…it will be better to…” this formulation sounds very strong. Both observational and modelling frameworks have their uses. I would present them as alternatives to one another, rather than stating that one is superior.
L84-86: it might be worth spending a few more words/sentences to introduce and define the concepts of AR1, variance, λAR1, and λVar, as they seem rather central to the presented results.
L90-93: “Motivated by the issue of what impacts of human activities especially forest management including PF and NF on vegetation resilience, this study used the latest global forest management dataset, multi-source satellite observations, and climate datasets to investigate. Here, this study mainly to:” unclear sentences.
L94: it might be worth introducing the “paired method” before mentioning it in the research questions.
L106: “In Spatial”, meaning?
L107: “decline” as per meaning become smaller, fewer, weaker, etc. (https://www.oxfordlearnersdictionaries.com/definition/english/decline_1) is an intransitive verb, meaning that it cannot have an object. In your formulation: “…human forest management activities have declined the resilience” you want to use a transitive verb such as “decreased”. Please check the use of decline vs decrease throughout the manuscript.
L108: Explain what the GHFC and GFMD are.
L111: Rationale for these metrics is needed. We have no explanation of what they do, how they are constructed and what assumptions hide behind. This needs to be clear in the main text, not guessed or ignored. Also, in the introduction you talked about four warning signals: AR1, variance, λAR1, and λVar. Why are the results based only on three, excluding the variance metric?
L112: is the starting time of the analyzed data 2000 or 2001? This seems inconsistent.
L112-113: “The results revealed…” what results? It is not clear to me what analyses were performed. Are the statements in this paragraph purely based on visual inspection, or is there more to it that is not mentioned? This should be clarified.
L117-118: “Forest fragments experienced higher levels of human disturbances and pressures, such as logging, clearing, or agriculture 38” It should be clarified that these are results from the literature, if so.
L123-124: “These… influence…” check consistency of singular/plural forms; “management…” practices?
L125-127: Again, not entirely clear if these statements are based solely on studies from the literature or if they simply agree with the literature.
Fig 1: confusing and not vizually clear. strangely structured. Axis titles are missing. Lat/lon indicators are missing. So far, LAI has never been mentioned in the text, so the subplots b and d come as a bit of a surprise. Background grid lines could aid the comparison of subplots b, d, e., and the x-axes could be aligned for more consistency and ease of interpretation. In the figure description: “as an example of a specific global grid point”, you could say what point you are presenting (e.g. central coordinates). Finally, the last part of the figure caption probably belongs to the main text: “The other examples in Central Africa… …found in Figure S2, showing similar spatial patterns”.
L149: it is here we discover there are paired grids. But how are they paired? some info needs to percolate in the main text
L149: “paired grids” again, it should be briefly mentioned before what this method is about.
L168-169: “over 40% of paired grids show higher resilience in PF than in NF” this is quite high! If this is the case, I wonder if the statements about management decreasing resilience is founded or whether you should be more conservative. Considering the relevant uncertainties, is the 60%-40% ratio really different than a 50%-50% ratio based on random noise?
L174-175: “P/PET (the Ratio of Precipitation to Potential Evapotranspiration)” it is good that you define P/PET, but you should maybe also mention what it is since it is the first time you mention it in the text, i.e. an aridity index.
L179: “these 15 resilience calculations…” not entirely clear what “these” refers to.
L183: “Fig. S7”, what happened to figure S6?
Figure 2: the color palettes chosen to represent the resilience difference and the datasets are very similar (browns and greens especially). This makes the interpretation of the figure difficult, so the colors should be differentiated more. Alternatively, subplots b, d, and f should come with their own color legend. In the figure description (here and in other figures) you mention a significance test, but it is not specified what was tested. The significant difference from 0?
L221-223: “since… which…” correct the structure of the sentence. Currently the main clause is missing.
L245: Mention what “background factors” are.
L249: Correct “focued” with “focused”.
L256: “explanatory” power?
L259: “other factors”, which ones? Maybe make a few examples.
Figure 4: I am a bit confused about SHAP values vs feature values. What do they represent and why do we need both? Are the feature values similar to effect coefficient? It should be made clearer even for readers that are not very familiar with the methodology.
L281-285: It is a bit strange that subplots e (PDSI) and i (temperature) are presented before a, b, c, etc.
L285: “…increases as the enhancement of…”, you mean “…increases with the increase of…”?
L287-288: “In this regard, results further suggests that PF are more sensitive to water resources than NF (Fig. 5e-5j)” this is a

bit hard to follow, since the results are presented via sensitivity difference. You could add a sentence to explain why this is the case.
L291: what happened to Figure S10?
L292: sensitivity to “soil water” or “soil moisture” might be more appropriate than “water” in general, as SM is what is shown in Fig. 5c).
L296-298: “Because the energy sensitivity difference between PF and NF gradually increases along with the wetter climate, and the variations of vegetation greenness is energy-limited in humid areas” missing a main clause.
L298-299: “As the enhancement of P/PET” does not seem grammatically correct. Change to “With increasing P/PET” or “As P/PET increases”.
Figure 5: Results descriptions and references to supplementary materials should probably go in the main text (results) and not in the figure description. Clarify what the significance test is testing. Also, “Vegetation greenness is limited by water in dry conditions and by energy in wet conditions, which may be related to shifts in resilience differences between PF and NF”: to me, this sounds related to actual resource limitations rather than resilience differences.
L327: λPF – λNF is a newly introduced metric/abbreviation. Explain what it is. It is also not found in the referenced figure 5.
L332: “…PF and NF are reducing in both…” should be “decreasing” instead of “reducing”.
L359: “However, a notable ambiguity arises from the definition of managed forests also came to this study” unclear sentence.
L388: “CEC” is not defined so far.
L399-401: repeated sentence.
L416-418: unclear sentence.
L420-422: “More interestingly, vegetation greenness is water-limited in dry regions (PF is more sensitive to water) and energy-limited in wet regions (NF is more sensitive to energy)”. To me, this sounds rather obvious. It is well known and intuitive that vegetation is limited by water where water is scarce, and energy becomes the limiting resource where water is abundant. Maybe you want to reformulate this to frame it more in the context of resilience and sensitivity.
L438: I believe one needs to reach this point to learn for the first time that these indices are made from LAI
L451: “decline” change to “decrease”.
L458: “the primary focus is on the differential impacts of forest management between PF on resilience” unclear sentence.
L480: “…and reduce uncertainties from land use change…” and following sentence: to me, this sounds more like land use change was ignored since you are only assuming that pixels do not change, but in reality, they might be changing. Am I missing something?
L491: “ASCE” define/explain what it means.
L503: change “photosynthetic” to “photosynthesis”.
L506: “PML” and “MODIS” you may want to explain the abbreviations or mention the source (which for MODIS you do a few lines below).
L511: this is a wrong use of kNDVI. see https://doi.org/10.1016/j.isprsjprs.2022.12.019
L516: “which are mainly from [MODIS]” and where are the other products from?
L518-519: “these two observed vegetation greenness or biomass are similar in a day” give a reference stating this.
L519: “But two times may be affected by the cloud” unclear sentence.
L524-526: You mention two contrasting spatial resolutions for the vegetation indices (500 m and 1 km). Which one is it?
L527: why are you limited to 1km by computational constraints? it seem very likely it could all be done at 500m
L530: you still have not explained how to calculate resilience. why not start with that before?
L531: If not already mentioned in the introduction, the “paired grids” concept should really be explained in detail here, as I am still wondering what it means.
L538: what is Simon et.al?
L541-543: mention supplementary figures explicitly.
L546: why monthly data? to do a robust time series analysis it is better to take all data available, not picking a single MVC value per month
L550: you need to explain what these are
L552: It is still unclear to me what the difference between AR1 and λAR1 is.
Formulas: try to keep the font of parameters consistent between the formulas and the main text.
L571: “same as the Xn” unclear.
L591-593: correct typing mistakes: “however”, “maxium”, “northern”.
L591: Explain what “MVC” stands for.
L592: Explain what “month datasets” are, or reformulate.
L593: “influencing” reformulate, unclear.
L595: “some CSD studies did not deal with missing values”, does that mean that missing values are excluded or kept as they are? Either way, it is not possible to “not deal with missing values”.
L603-605: “Not only because of the strong relationship between resilience difference between PF and NF and background variables, …” missing main clause. Also, this part specifically sounds as if you chose the method based on the results because “it looks good”, instead of basing the decision on hypotheses or research gaps.
L607: “widely utilized in ecosystem studies”, then maybe give a few more references instead of only one.
L612: what does “its” refer to?
L615” what does “their” refer to?
L621: “datasets availability section” should be corrected to “Data and code availability section”.
L671: “low probability pixels” what does this refer to, low probability of what?
L681-683: “GFMD is the result of a classification algorithm that takes in input various factors including vegetation data, the same data to calculate resilience. It may undermine the independence between forest management and resilience”. This sounds like a potentially huge issue. If the input data are not independent, how can you say anything about the results? The following parts were you control with independent data should really be mentioned more centrally.
L702-703: “This is aligned with our main conclusion. The above analysis results can be seen in the discussion and

supplementary section" good, as mentioned above, I think this confirmation should be more central to the results. Also, I could not find it in the discussion, where it should really be, rather than hidden at the end of the methods or in the supplementary materials.
L707: the data here means the data produced by the paper, not just the data used as input (although the latter is also needed)

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)
I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
This study explores the resilience of forests under varying management intensities, ranging from unmanaged natural forests to managed, planted forests and plantations. The research examines how resilience, as defined by the critical slowing down (CSD) theory, differs among forests with different levels of management. The study concludes that natural, unmanaged forests exhibit greater resilience, while planted forests are more vulnerable and thus show lower resilience. Additionally, the study investigates the factors influencing these differences, reporting that wetter regions are less vulnerable and, consequently, more suitable for planted forests.
I find the study both interesting and novel, contributing important findings beyond previous similar studies. However, several points need to be addressed to strengthen the results. The authors conducted multiple sensitivity analyses to demonstrate that the findings are consistent, considering factors such as the sliding window width, forest coverage threshold, definitions of planted (PF) and natural forests (NF), reliability of forest management maps, and preprocessing methods of satellite vegetation data. However, one crucial aspect that should be considered is forest age: natural and managed forests are typically older than planted forests and plantations, and resilience is known to vary among different age groups. Therefore, I recommend including an additional sensitivity analysis to compare planted forests with younger unmanaged stands, to verify that the results are not merely due to age differences between PF and NF. Furthermore, in the final chapter discussing potential limitations, the authors should expand on the limitations of the CSD theory.
Aside from these points, I have only a few minor comments detailed below:
Abstract: The abstract lacks specific mention of the aspect of resilience studied. I suggest mentioning the CSD theory in the abstract.
Introduction: The introduction is well-prepared, effectively outlining the concept of resilience, identifying the research gap, and clearly stating the study's objectives.
Methods: The authors utilized validated datasets, such as GFMD, the TerraClimate database, and MODIS. Although the methods involve some subjective choices, the sensitivity analysis provides reassurance about the robustness of the results. The classification of resilience into stable, increasing, or decreasing is based on a well-established theory and is therefore justified.

Detailed Comments:
Line 106: The phrase "In Spatial" is unclear—please clarify.
Line 108: Introduce the abbreviations GHFC and GFMD here (or in the Introduction) as they are used for the first time.
Line 176: Could you provide an explanation for this pattern already here? Why are the results different in these regions?
Line 183: Please check for grammatical accuracy—consider changing to "showing similar patterns" without "the."
Line 242: Instead of "cycle," did you mean "circle"?
Line 320: Please provide references to the "previous studies" that also documented a global decline in resilience.
Line 425: I suggest adding a sensitivity analysis comparing younger NF and younger PN here. Additionally, consider mentioning the limitations of the CSD theory.
Line 538: The reference "Simon et al." is missing from the reference list.
Lines 614-615: Why didn't you use age as a predictor variable? Age could significantly impact resilience and potentially affect your findings. Please elaborate on this. Since PF are generally younger than NF, I recommend an additional sensitivity analysis comparing PF and NF of similar ages.
Lines 617-621: Although you mention that "soil and vegetation variables" are detailed in the dataset availability section, I believe these should be briefly described in the Methods section, as they are critical inputs for your model.
Line 650: I couldn't find the reference to "Taylor et al."—please ensure it's included in the reference list.

(Remarks on code availability)
I have not reviewed the code, but I have confirmed that it is available online as stated by the authors.

Reviewer #4

(Remarks to the Author)

Review report
The study is focused on the forest management practices (human factors) and the (natural factors) driving forest ecosystem resilience. There are sustainable forest management practices that support forest ecosystem resilience while other unsustainable practices undermine forest ecosystem resilience. This is a very interesting topic, especially as the resilience of forest ecosystems is critical to the supply if ecosystem services benefiting people. The effect of natural factors that also influence forest ecosystem resilience, and that is a key objective of the manuscript is not capture in the manuscript title? The resilience of forest ecosystems is not only explained by forest management practices (type and intensity) but also more so, by biophysical factors, including climate change, and as such should be captured in the title.
Besides, I have some general major concerns.
Comparing planted forests and natural forests seems odd for me, as the difference in their resilience is quite obvious from their management practices. The results stated in the abstract are quite obvious: For example, “forest resilience suffers greater losses under intensive management, as seen in human plantations like agroforestry, which are much less resilient than unmanaged natural forests”. And “Such resilience difference between PF and NF is related to background conditions such as local climate, vegetation, and soil”. Topography also plays a role? These are obvious well documented results and do not provide any novelty, even where a quantified difference is provided. Is the resilience of well managed planted forests at the same level as or even more than unmanaged natural forests?
My suggestion is to compare similar forest ecosystem types (e.g., managed and unmanaged natural forests; planted commercial forests and planted forests mainly aimed to restore ecosystem functions and service, etc.). Comparing similar forest ecosystem types with different or similar forest management practices and biophysical conditions would be more intriguing in understanding differences in their resilience and inform restoration and conservation decision making.
Comparing within planted forests with primary objective for restoration functions and services and planted forests under the commercial forestry approach can enhance our understanding of the resilience of these ecosystems in terms of supply of ecosystem services. The resilience of planted forests monocultures and multiple tree species forests can be compared to further understand why forest monocultures are discouraged.
Comparing managed and unmanaged natural forests can provided important insight on the effect of forest silvicultural management practices on forest ecosystem resilience. Unmanaged old growth forest may sequester less carbon due to older age trees compared to young secondary regrowth forests. Do unmanaged natural forests need some sought of management intervention to increase resilience or should they be kept intact?
The resilience of forest ecosystems is not only explained by forest management practices (type and intensity) but also more so, by biophysical factors. But it is clear this was not investigated in this study, making the scope so one dimensional. This is clearly acknowledged by the authors. For example:
In line 446 to 448, the authors clearly stated that “…the narrowing disparity in resilience between PF and NF may be attributed to multiple factors, including vegetation physiology, climate change, and the intensity of human activities. This complex issue merits further investigation”.
In lines 459 to 460, “… but there is less exploration of whether the climate, soil, and vegetation background can mitigate the effects of other forest managements”.

(Remarks on code availability)

Version 1:

Reviewer comments:

Reviewer #1

(Remarks to the Author)

Overall impression

The authors have made a remarkable effort to make the manuscript more legible. They have effectively answered several concerns of the reviewers. However, now that it is clearer (notably on the metrics used), my overall feeling is that there is inherent flaws in this study. Unfortunately, these flaws confirm to me the first impression that it should not be accepted. I will elaborate below my main concern, and the expose some of the more specific issues I encountered while reading the document.

General issues

The main problem I have with this study relates to the use of the CSD theory specifically on these global vegetation time series of remote-sensing derived indicators in view of characterizing alleged differences in resilience between natural and

planted forests (i.e. the main goal of the study). As the authors state (L85) "According to CSD theory, as systems approach tipping points where system states
change, their sensitivity to disturbances increases, and recovery from disturbances slows down, indicating a loss of resilience". To begin with, it would be necessary to state specifically how these concepts are specifically mapped to the study case at hand of forests. First, what are the tipping points we are considering? Is it "collapse" of the forest? how is that defined specifically for PF and NF? Defining the tipping point seems necessary to me so as to say what signal we should be using to monitor the disturbances for establishing the characterization of resilience. Then, the next question is: what specifically are we considering as "disturbances"? Windthrows? Fires? bark beetle infestations? clear-cuts? Later in the discussion it states (L500) "abrupt events and human disturbances lead to non-stationary conditions" and that this limits the accuracy. So I do not exactly understand what disturbances we are trying to see in the signal. This adds questions then to whether the signal used are appropriate in terms of (1) what they measure, (2) over what spatial supports, and (3) with what temporal frequency. For point (1), all metrics used fundamentally relate to temporal variations in greenness (whether they be LAI, kNDVI, fAPAR, GPP, they are reflect mainly greenness and the information comes mostly from NIR and Red reflectances). Are changes in greenness over 5 by 5km (point (2)) detectable with a frequency of one or twice a month (point (3)) really capturing the target "disturbances" that are not man-made and not abrupt? Is this the approapriate tool for this question?

To bring a bit more of a specific detail of why this is problematic for this question, consider that managed forest will be more subjected to clear cuts by definition. These will entail stronger drops in greenness than for natural forests, so when this is averaged of coarse spatial resolutino of 5km, the mixed signal will be more strongly affected (on average) for the managed forest than natural forest. First, is any normalization used here? will the AR1 metric not be artificially affected by this? Then, the recovery could be faster in a planted forest if fast-growing species are planted (and assisted, with protection from herbivory for instance) than if natural regeneration occurs. This would be reflected in the metric by suggesting managed forest would have higher resilience (as faster recovery), and yet this is counter to what the results suggest. If such clear-cuts events should not be considered (as suggested by some phrases in the text cited above), then how is this prevented from affecting the results and the conclusions?

I can go on with more concerns, such as the questioning of whether there are enought samples in time to have sufficient frequency to sample the stable state as suggested in Figure 2c and 2d of Scheffer et al (2015), mentioned by the authors as the reference. With monthly data over 20 years, I have some serieous doubts that this can apply so seamlessly. I would think you would then need daily data from satellite, which is of course available, but then you need to deal with changes in viewed angles and missing data, which I suppose incapacitate the use of these AR techniques. Furthermore, the MVC technique to composite means the actual time stamps of each month is not regularly separated from the former or from the next, so this further means the AR techique is applied in sup-optimal conditions (i.e. non regular timesteps). All these issues are not addressed.

I know others have used these metrics on such similar data before (e.g. Forzieri et al.), but that does not convince me either, as I believe the same flaw applies there. I would require a proper rebuttal plus explanation to address these specific issues to accept the premises of this manuscript. Below I further provide more (non exhaustive) specific points to help the authors improve the text.

Specific points

L73: Management should not be plural. Perhaps change to "types of management"
L82: CSD should be spelt out here
L91: These new metrics should be explained, even if briefly, not just cited. The reader should get here the diffrence between AR1 and lambda_AR1. What does the lambda stand for? What is Lambda variance?
L94: what is "variance" by itself? you need to be more specific... the temporal variance of the signal... and over how long of a period?
L97: In this paragraph you speak of statistical metrics, but you should have also mentioned the signal on which the metrics need to be applied, otherwise it is pointless. They cannot be applied to a signal that does not speak about resilience. So first mention how RS signals can effectiveely (or not) be sensitive to properties that speak to resilience, and only then introduce the statistical metrics.
L104: Not clear why ML comes into play? Why would it be needed? needs to be introduced. Also the "impacts of biophysical factors". What are these "biophysical factors"?
L110: "biophysical factors" should have been mentioned and described before

L117: why tree cover from 2000 and management for 2015? the former is available for 2000
L122: kNDVI is with lowercase k
L127: again, the acronyms are meaningless and useless if we do not know what they measure

L499: Here is a line that puzzles me. If we need stationary behaviour, and "abrupt events and human disturbances lead to non-stationary conditions"

L502: This scale dependence is crucial to take into account in my opinion as in this case, natural forest will likely have different (smaller) spatial extents of perturbations than managed forest (where, by definition, there would be more clear cuts).

L508: GPP is not a biomass index. And the way it is calculated (from MODIS) is very likely affected by the same saturation effect of VIs.

L614: MVC only makes sense for NDVI, not for (downstream) other products (e.g. LAI, fAPAR, GPP, etc.), so be precise on whether you use it only for NDVI or not, and what are you using for the other signals. Also, the reference of MVC should be the original, Holben et at.

L668: doesn't this window width also depend on the frequency of observations?

L672: Why MVC?

L674: you should be more specific, as here what you are doing (with MVC), is to downsample the frequency of actual observations. if the monthly values are effectively a single value in the time series (as with MVC), the time between consecutive observations is not a month, but van vary. So what does that do for CSD?

L690: How good is the Besnard et al data to differentiate forest age data over a 0.25 by 0.25dd area? it is a requirement that this is very high if this is used in a moving window approach, otherwise this does not provide robust info

The following are 3 other specific points regarding the response to reviewers.

1
Figure 1. Action taken. (2) Background grid lines have also been added to subplots b, d, and e to aid comparison (add the Lat/Lon in the figures).

I do not see any background grid lines. Maybe they were lost due to the low-resolution of the figure?

2
L168-169: “over 40% of paired grids show higher resilience in PF than in NF” this is quite high! If this is the case, I wonder if the statements about management decreasing resilience is founded or whether you should be more conservative. Considering the relevant uncertainties, is the 60%-40% ratio really different than a 50%-50% ratio based on random noise?

R: Thank you for this important question. Our intention is not only to highlight that certain management activities can diminish resilience, but also to emphasize the complexity and variability of these effects. While forest management activities associated with high human pressure—such as intensive plantations—can significantly reduce forest resilience, there are also management practices characterized by lower anthropogenic pressure that may enhance resilience under specific conditions. For instance, as illustrated in Figure 2, the most intensive form of human planting (i.e., plantations) shows the greatest negative impact on resilience, whereas less intensive afforestation practices (i.e., planted forests) have a relatively weaker effect on reducing resilience. Action taken. Nonetheless, we recognize that the reader may be misled here, so we have rephrased the paragraph topic, now reading as (Lines 184~186). “Although forest management activities with strong human pressures have weakened forest resilience, managements with lower anthropogenic pressures can enhance resilience under specific conditions (Fig. 2).”

Not sure if this is answered. My original comment referred to the fact that the percentage of pixels where one or the other conclusion could be drawn is potentially indistinguishable from random noise. The authors state that “40% of paired grids show higher resilience in PF than in NF” (L186-188). This means that the main outcome is that 60% of paired grids show lower resilience in PF than in NF. Given the many confounding factors affecting natural systems, how confident are you that these results are different than a null hypothesis where 50% of paired grids show higher resilience in PF than in NF, and 50% show lower resilience in PF than NF? While I agree with the explanation provided by the authors about the management effects on forests, this misses to answer the raised concern about the percentages and random noise, which is a statistical issue.

3
Supplementary Figures. Given the updated figure specifications (e.g. color scales), try to also update the supplementary figures to be consistent with the figures in the main text. E.g.: the maps in Figure 2 and Figure S4 have now a different color bar.

(Remarks on code availability)

Reviewer #2

(Remarks to the Author)

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

(Remarks on code availability)

Reviewer #3

(Remarks to the Author)
The authors have adequately addressed my primary concern regarding the effect of age on the main conclusions of their study. Additionally, they have satisfactorily resolved other issues raised in my initial review. After carefully rereading the revised manuscript, I find no further comments or concerns.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
Comments for authors
The focus of the study on forest management types, biophysical factors and management practices, and forest ecosystem resilience" is timely and interesting, as the results will help to inform decision-making on forest management best practices for various forest management types to build resilient forest ecosystems in a changing climate.
I took the time to review the author response to all the reviewers' comments, supplementary materials and main document, and I think the authors have greatly improved on the manuscript in terms of technical and scientific methodology and clarity in the presentation of the results. In particular, the classification of the forest management types by the authors is much more improved and clearer, enabling better comparison between forest management types. The conceptual figure 2(I) that illustrates the decreasing vegetation resilience across the different forest management types due to human impact is much appreciated. The objectives of the study are clearer and are well answered.
The role of forest management practices (e.g., logging rotation cycle, restoration efforts - for managed natural and planted forests) and some biophysical factors in shaping the resilience of forest types/ecosystems has been addressed in greater detail in the text and figures, which has significantly strengthen the ecological context and applicability of the findings. Furthermore, the authors have provided clearer explanations of the models and their validation procedures, and essentially specified how the models were assessed, including performance metrics, validation techniques, and potential sources of bias and uncertainty in the spatial data. Therefore, I agree with the methods used and results reached by the authors. However, I still have some major comments/suggestions
Major comments
I think that the results are mainly confirmatory of exiting factual knowledge and merely qualitative in the way that they are presented in the text. I think that the results could be better quantified, precisely when comparing the resilience between forest management types. In particular, instead of using "higher" to qualify unmanaged natural forest's superior resilience over managed forests and planted forest, for example, the authors could quantify the resilience difference in terms of how much divergence or variation that the study found (resilience difference) that may differ from previous global studies. For example, how many times or by how many folds is one forest management type more resilient than the other? This is an important quantification information that provides more context for comparison than merely using the words "higher" or "greater" and "lower" resilience to express the resilience difference between forest management types or forest ecosystems. For example, if an unmanaged natural forest is subject to intensive management practices, how many folds will be the decrease in the resilience factor? This is because, it is factual knowledge that unmanaged NF are older and intact, and thus, are more resilient than managed NF and PF in all climatic and biogeographical environments, and what this means in the context of climate change. Thus, providing the magnitude of the resilience difference in terms of numbers (number of times or folds resilience increased or decreased between forest management types due to forest management practices) should be one of the research gaps to be filled in the existing literature. As a demonstration, in forest carbon management studies that compare natural and regenerating forests, for example, results are presented in a way that is more captivating such as "old growth (unmanaged) natural forest sequester 10 to 20 times more carbon than young secondary or regenerating forest (managed forest)"; is such quantitative comparison possible with the current study between forest management types or forest ecosystems?
Minor comments
The title is still unclear in terms of the structural presentation of the independent and dependent variables and the interrelationship. The biophysical factors and forest management practices determine the resilience of forest management types. The focus of the study is on the resilience differences of forest management types as influenced by biophysical factors and forest management practices? I suggest that the authors should provide a more thoughtful and informative title.
In the abstract, natural forest should be used instead of native forest, as this is the terminology used throughout the manuscript; the efficient "supply" or "provision" (choose one term) of ecosystem services should be used instead of "functioning"; external disturbances (i.e. ecosystem resilience) is unclear, since external disturbance does not mean the same thing as ecosystem resilience; the former influences the latter.
The authors should categorically state in the abstract that unmanaged natural forest demonstrated higher (quantified in terms of magnitude of resilience difference) resilience than managed natural and planted forests (or all the other forest types). The authors clearly stated this result in the text as follows "… the results by the nearest neighbor method showed that unmanaged NF demonstrate higher resilience than managed", but not in the abstract. This is the key result and should be

stated in the abstract first, before stating how these may vary under different biophysical conditions such as wet climate (water availability/deficiency modulating the resilience difference effect of forest management types). I think that unmanaged natural forest is more resilient under all climate conditions, though I agree that adapted best management practices can improve resilience under favourable biophysical conditions.
In the abstract, "These differences are influenced by biophysical contexts such as local climate, vegetation, and soil". However, I think the authors should add one sentence on the limitation of the study, especially in terms of the limited variables included in the model in the abstract.
The full stop (.) mark comes after the quotation (") mark.
There are too many abbreviations and acronyms in the text, please reduce them to a minimum to improve on the readability of the manuscript. For example, all the forest management types could be written in full in the text.
There are still some issues with grammatical accuracy that need greater attention and revision before publication. The authors can seek the help of a native English speaker colleague in the same field of research.
Figure S1. A clearer font (without text effects) should be used to label the forest management types, a single font should be used for all labels in all the figures.
Figure S12. The shadows in the labels or lettering should be remove.
Figure S22. The inclination (e.g., slant vertical) of the letterings of forest management types below the figure should be improved for better readability.
Figure S27. The labels of forest management types below the figure should be restricted to fall outside the figure.
Does differences in species composition between forest ecosystems influenced the ecosystem resilience in this study? Was this factor considered as species richness, evenness and diversity under the vegetation variable, in addition to AGB variable? At least, using more vegetation variables in the model would help in detecting the actual influence of biophysical factors on forest ecosystem resilience.
"This discrepancy is potentially due to the imbalance of samples in the crowdsourced points, where samples of forest managements like PF are less as compared to NF" – Please provide precise numbers in the manuscript text instead of using "less" and "more" qualifiers.

(Remarks on code availability)

Version 2:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
Attached file.

(Remarks on code availability)

Reviewer #4

(Remarks to the Author)
The authors have provided a comprehensive explanation of the limitations of the study, in terms of variable input, analysis and future directions. They also have responded satisfactorily to my comments, especially presenting the magnitude of difference in the resilience between forest management types. Therefore, I approve of their revision but encourage the authors to consider highlighting the limitations of the study on the remote sensing and disturbance regime effect sides.

The limitations on the remote sensing side in influencing the data analysis and results are profound, but can be overlooked due to the global scale nature of the study. The spectral signatures of vegetation types and management conditions undergo constant change in strength and amount, following various natural and human disturbances in forest ecosystems, whether continuous, abrupt or subtle disturbances. So this in itself creates challenges for remote sensing data processing and analysis, in addition to image distortion challenges. So, I think the results offer a foundation for future studies using more fine resolution data such as from the global ecosystem dynamics investigation (GEDI). I encourage the authors to add these remote sensing related limitations in the manuscript to inform the readers of their awareness and transparency.

The limitations on the effect of specific disturbance type on the resilience of forest management types is also glaring, but dwelling on this in this study is unfeasible (it is a topic on its own), as such should be highlighted in the discussion and added as a recommendation for future studies. Different disturbance regimes affect different forest ecosystems differently in spatial and temporal scale.

(Remarks on code availability)

Version 3:

Reviewer comments:

Reviewer #2

(Remarks to the Author)
The authors have satisfactorily responded to all my main comments. Before acceptance, I would only suggest the following minor changes:

Supplementary L318-319: "PF generally possess lower structural complexity and shallower, simplified root systems", please add source(s) for this statement.

Supplementary Figures S22, S28...: you added the coefficient of variation, but labelled this as R instead of R2. If I am not mistaken, this should be R^2 (R-squared). Please check this and double check that it is kept consistent across all figures. Similarly, the p-values are missing from e.g., Figures S28-S29.

(Remarks on code availability)

Response to Reviewers

The authors sincerely thank the anonymous reviewers for their valuable and constructive feedback. We have carefully addressed all concerns and revised the manuscript accordingly. Below, we provide detailed responses to each comment from the Editor and reviewers. In the revised manuscript, reviewer comments are shown in black, our revisions are highlighted in red, and our responses are indicated in blue. Please note that all line numbers refer to the revised version of the manuscript. Additionally, the references cited within the red-colored text are available in the tracked manuscript, while those in the blue responses are included in this document.

Response to Reviewer #1

Reviewer#1

Overall impression

The manuscript attempts an analysis of the sensitivity of forest resilience to human disturbance and environmental drivers in different forest management categories, arguing that natural forests have higher resilience than managed forests in 60% of the cases. While the subject is of great interest, the manuscript is unfortunately not convincing. The techniques used are inappropriately explained and the limitations of the underlying data are not well exposed. Furthermore, the overall quality and style of the manuscript is not in-line with what is expected in such a journal. References are incomplete, lacking journal names, sometimes pages. There are quite some errors in language that should have been picked up by any spelling corrector. Some phrases seem unfinished (no verb, or no object, see L93). Then technically too, various crucial methodological details are not explained in the main text, making understanding and following the text quite difficult, and undermining the credibility of the claims that the manuscript tries to put forward. Given these various points, the overall recommendation would be rejection. However, if the authors want to improve the manuscript, hopefully this review can help them. Below is a list of more general issues followed a more detailed list of specific points.

R: We appreciate your review and constructive feedback on our manuscript. Your comments have provided useful insights that have guided us in revising the manuscript. We have addressed all your specific concerns as detailed in the following lines. We have incorporated additional analyses to investigate the impact of varying spatial resolutions (i.e., 500m or 1000m) on our findings. The Introduction and *Method Section* have been substantially revised to better differentiate between multiple indices within the CSD framework, with particular emphasis on justifying our focus on the λAR1, λVar, and AR1. In addition, we added a discussion of sensitivity analysis based on crowdsourced points related to the independence of GFMD and LAI data. The abstract has been restructured to more prominently highlight resilience quantification. We have thoroughly reviewed the reference format and added missing journal names, volume numbers, and page numbers. The detailed reversion can be found in the response and tracked manuscript.

General issues

1. - The abstract gives no clue on the nature of the work. How is resilience quantified? Model simulations? satellite metrics? ground-data?

R: We appreciate your valuable comments regarding the clue of the abstract and its clarity. Resilience is quantified based on the Critical Slowing Down (CSD) theory, which posits that as a system approaches a tipping point, its recovery rate from perturbations slows down. Forest resilience is captured through three resilience metrics (AR1, λAR1, and λVar) derived from satellite-based vegetation indices (LAI, FPAR, KNDVI, GPP_MODIS, GPP_PML).

Action Taken. In response, we have made the following revisions to address your concerns: (1) Descriptions of CSD Theory, (2) Clarification of Resilience Quantification, and (3) Clarification of method and datasets from satellite-based metrics in the abstract. Here is the modified Abstract (Lines 4~22):

"Policymakers are concerned that newly planted trees may not provide ecosystem services as effectively as native forests. The efficient functioning of ecosystem services relies on the capacity of forest ecosystems to recover from external disturbances (i.e., ecosystem resilience). This study utilizes the critical slowing down (CSD) theory to quantify resilience differences among forest management types, particularly between planted forests (PF) and natural forests (NF). CSD resilience indicators such as AR1, λAR1, and λVar are estimated from satellite-based vegetation indices. This study further investigates the biophysical drivers for the resilience difference between PF and NF by machine learning. Results show that most forest management activities weaken forest resilience. Forest resilience experiences greater losses under intensive managements However, some forest managements can enhance resilience under specific conditions, such as wet climates. PF display higher resilience than NF in P/PET (the ratio of precipitation to potential evapotranspiration) > 1.5 (wetter climates), while NF is more resilient in P/PET < 1.5 (drier climates). We further detect such resilience differences in all climate zones between PF and NF decreasing from 2001 to 2015. These differences are influenced by biophysical contexts such as local climate, vegetation, and soil. Regions characterized by cold, wet conditions, dense vegetation, and higher soil fertility minimize differences in resilience, suggesting that rational site selection for afforestation can enhance forest resilience compared to native forests."

2. - Consider that the methodology section comes at the end, so any concept, method, dataset, or abbreviation needs to be introduced before being dropped on the reader without context. A few examples are listed in the detailed comments.

R: We agree that concepts, methods, datasets, and abbreviations should be introduced before being mentioned in the text. **Action taken.** According to the reviewer's comment, we have introduced key methodological aspects throughout the manuscript and better contextualized their rationale before presenting and interpreting results. Here are the modifications:

(1) Introduction and explanation of paired grid method

We have introduced and explained the concept and methodology of paired grids in both the Introduction and *Result Section* to ensure readers understand the approach:

"To differentiate the impacts of human activities from natural variability, we employ a paired method that compares geographically proximate forest areas under different management practices." (In the Introductions, Lines 102~104)

"Specifically, grids are paired using the sliding window method, which moves a fixed-size window across the study area. Within each window, paired grids are selected based on similar forest age (differences of ≤10 years), elevation (differences of ≤500 meters), and a minimum of 20% forest coverage for each management type (more details in the Method Section). We point out that the effects of potential covariates have been ruled out within the sliding window approach." (In the Results, Lines 159~165)

(2) Clarification of abbreviations and datasets

We have provided the full interpretations of all abbreviations such as variables involved in the

XGBT model:

"Abbreviations of biophysical factors include: Temp-temperature, P/PET-the ratio of precipitation to potential evapotranspiration, AGB-above-ground biomass, BGB-below-ground biomass, CEC-cation exchange capacity, RD-root depth, Sand-sand fraction, and Clay-clay fraction." (In the caption of Figure 4, Lines 304~307)

"The eight environmental factors include annual temperature (Temp derived from TeerraClimate), dryness index (P/PET derived from TeerraClimate), root depth (RD[92]), above-ground biomass (AGB[93]), below-ground biomass (BGB[93]), clay fraction (Clay[94]), sand fraction (Sand[94]), and cation exchange capacity (CEC[94]). These factors are selected for their known influence on forest resilience, including both climatic and land-surface variables." (In Method, Lines 735~740)

We ensured that all dataset abbreviations were introduced with their full names upon first mentioning. Additionally, a full-text check was conducted to confirm the proper introduction of all abbreviations.

"Time series of multiple vegetation indices were retrieved for the selected forest pixels over the period 2001-2015, including leaf area index (LAI), kernel normalized difference vegetation index (KNDVI), fraction of absorbed photosynthetically active radiation (FAPAR), gross primary productivity from MODIS (GPP_MODIS), and GPP from the Penman-Monteith-Leuning model (GPP_PML)." (Lines 120~124)

"We derived the reference distribution of forest managements for the year 2000 (Fig. 1a) by overlapping tree cover data available for the year 2000 derived from the Global Hansen Forest Change Map (GHFC) and the spatial mapping of forest management types retrieved from the Global Forest Management Datasets (GFMD) for the year 2015." (Lines 116~119)

We mentioned the abbreviations of different management forests in the caption of Figure 2. In addition, the details of different management forests are shown in the *Method Section* and the supplementary.

"(k) Further comparisons of resilience are performed across various forest management types using the same methodologies. These comparisons include NF vs. PF, NF vs. Plantations, NF WOM vs. F WM, NF WOM vs. NF WM, and NF WOM vs. Plantations. The notation "X vs. Y" indicates the resilience difference, calculated as the resilience of Y minus that of X. Statistical significance tests are performed to assess whether these resilience differences are statistically distinguishable from zero. Bars marked with an asterisk (*) denote statistical significance at $p < 0.05$, while bars labeled with 'NS' denote non-significant differences. (l) A conceptual figure illustrates different forest management types (e.g., NF WOM). Abbreviations used in the figure include: NF-Natural Forests, PF-Planted Forests, NF WOM-Natural Forests without Managements, NF WM-Natural Forests with Managements, and F WM-Forests with Managements." (In the caption of Figure 2, Lines 206~216)

(3) Extending resilience metrics concept

We extended the introduction and quantification of resilience metrics, clarifying the differences among them, particularly emphasizing why the metric of variance is less used. Changes in the revised text are in Lines 88~97 and Lines 621~641.

3. - In the introduction, four warning signals are presented: AR1, variance, λAR1, and λVar. Yet, the results are based only on three, excluding the variance metric. This needs to be motivated or risk feeling like cherry-picking.

R: We agree with the reviewer's comments. It needs to be clear explanation of excluding the variance metric. In the revised manuscript, we have provided a clear justification. Although variance serves as a warning signal for potential losses in resilience or stability, its application is limited for several reasons: **(1) Sensitivity to Environmental Noise**: Variance is highly influenced by external environmental factors (i.e., climate change or human activities), which can introduce noise and obscure the true system dynamics, particularly in observational datasets[1]. **(2) Lower Sensitivity to Critical Transitions**: Variance is less effective in detecting critical transitions compared to AR1-based metrics, as it does not capture recovery dynamics as robustly[1]. A recent study showed that AR1 increased, while the variance decreased for global vegetation resilience[2] (the signs of variance showed the resilience increasing at a global scale) (Fig. R1). Many other studies have reported that global vegetation resilience has declined in three decades[2-4]. **(3) Physics interpretation:** Recent studies have increasingly focused on the AR1, λAR1, and λVar due to their closer alignment with the dynamic concept of resilience [4-6]. Whereas the concept of variance is more closely related to the fluctuation of the system. **(4) Previous studies:** Recently many studies have begun to focus on the application of λAR1 and λVar[4,7,8], due to their improvement from variance and AR1. According to the formula, the newly proposed λVar calculation process includes not only variance but also AR1, which may be a more comprehensive index. λAR1 is derived by applying a logarithmic transformation to AR1, giving it a more direct physical interpretation related to recovery rates.

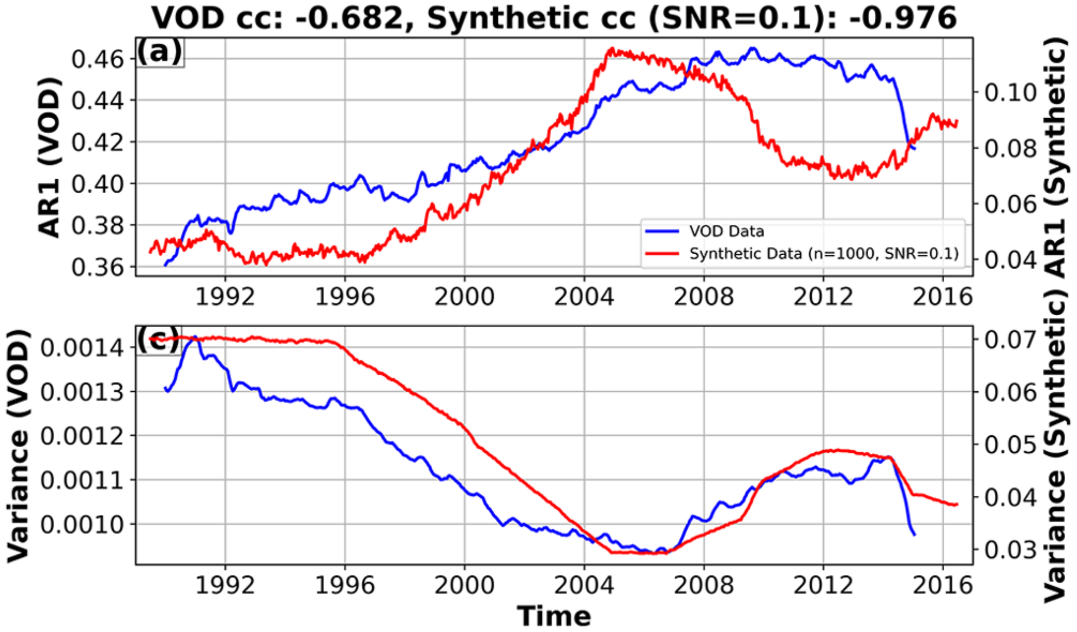


Figure R1. (a) Median AR1 for vegetation optical depth (VOD) data (blue; median taken over all AR1 series globally). **(c)** Same as **(a)** but for the variance. The blue line is the VOD datasets from the satellite, while the red line indicates results generated by satellite-based data with random noise (Signal-to-Noise Ratio (SNR) = 0.1). This figure is derived from a previous study[2].

Smith et.al applied the λAR1 and λVar, which indicate the rate of recovery from disturbances and provide deeper insights into system resilience[4]. AR1 (autocorrelation) indicators mainly capture a system's "memory effects" or delays in returning to equilibrium, while variance focuses on the overall magnitude of deviations from the mean[1]. In contrast, λAR1 and λVar quantify the post-disturbance recovery rate, directly characterizing how quickly the system returns from a non-equilibrium state to a stable equilibrium[4]. In other words, they provide not just a statistical proxy for resilience, but a measure more closely aligned with its true ecological or dynamical meaning (the recovery rate after disturbance)[3]. Furthermore, variance is easily influenced by external environmental noise, making it challenging to distinguish genuine critical transition signals from random fluctuations[9]. By directly indicating recovery rates, λAR1 and λVar help isolate system-intrinsic dynamics and true recovery properties, rather than being confounded by noise-induced temporal correlations and variance changes[7]. While an increase in AR1 is often regarded as an early warning signal of an impending critical transition, studies have shown that the variance indicator may not display similarly pronounced changes in certain contexts[2]. The direct measurement of recovery rates via λAR1 and λVar allows for the detection of subtle declines in resilience more sensitively, providing clearer and more robust evidence of resilience deterioration even when variance does not exhibit a noticeable increase or decrease[7]. Additionally, our analysis based on λVar has demonstrated similar patterns to results by λAR1, indicating that the choice of indicators does not significantly affect our results. Furthermore, from a formula perspective, λVar incorporates both AR1 and variance information in the quantification process, making it a more comprehensive index. In summary, our study has excluded variance metric, instead focusing on AR1, λAR1, and λVar—with our analysis by the three metrics showing similar patterns (see Supplementary Fig. S3-S6 and Supplementary Fig. S17–S21).

Action taken. In response to your comments, we modified some texts in the Introduction to clarify the reasons for using three metrics but exclude the variance (briefly introduce). Here are the modified texts: (Lines 88~97)

"The AR1 measures how strongly the current state of a system is influenced by its previous state[18], and the variance measures the magnitude of fluctuations around the system's equilibrium[43]. In addition, new metrics such as λAR1 and λVar have been developed, which are more closely related to the rate of vegetation recovery from disturbances, further aligning with the concept of resilience in dynamic systems[19, 44, 45]. Compared to the other three metrics, variance is susceptibly affected by environmental noise, and thus is challenging to detect critical transitions. As a result, recent studies have increasingly focused on AR1, λAR1, and λVar due to their stronger alignment with the concept of forest resilience[45, 46, 47]."

Moreover, we have specifically introduced and explained it in the *Method Section*. (Lines 621~637)

"Based on CSD theory[16], various indicators, including lag-1 autoregression (AR1) and variance, λAR1, and λVar, have been widely employed to quantify ecosystem resilience[19, 44]. However, compared to other indicators, variance is highly susceptible to environmental noise, leading to less reliability in detecting critical tipping points[43]. For instance, while global vegetation resilience is reported to be in decline, as reflected by significantly increasing AR1 values and abrupt disturbances[18], variance has instead shown a decreasing trajectory[68]. As a result, recent studies have increasingly favored AR1 as an alternative to variance, as it can detect resilience and its changes across various ecosystems, including forests, wetlands, and grasslands, before regime shifts occur[39, [44, 45]. Additionally, in light of the influence of observational space-temporal resolution and observational noise, complementary indicators such as λAR1 and λVar have been introduced[45]. These metrics quantify recovery rates, measuring how quickly the system returns from a non-equilibrium state to a stable equilibrium[85]. In other words, they provide not only a statistical proxy for resilience but also a measure more closely aligned with its true ecological state. Consequently, we evaluated forest resilience using AR1, λAR1, and λVar."

4.- Check that the references to the figures are correct throughout the document. It felt that some were not, which made following along sometimes very confusing. Also, Figures and supplementary Figures should be always presented in order, and all should be referenced in the text.

R: **Action taken.** According to the reviewer's comment, we have made the following improvements in references: (1) Figure and Supplementary Figure References: We have ensured that all figures and supplementary figures are correctly referenced in the text and are presented in the proper order. Any discrepancies or missing references have been corrected, ensuring clarity and consistency throughout the document (see *References Section*). (2) References in the Reference List: We have checked the entire reference list for completeness and accuracy. Missing details, such as journal names, have been added.

5.- Statements based on studies from the literature should probably be moved all to the discussion, as in the results it is not always clear if the statements made are based on own results or the literature. Alternatively, it should be made very clear whenever a reference is cited as basis for a statement, or as support of own conclusions.

R: We agree that it is essential to distinguish between statements derived from our results and those based on the literature. **Action taken.** To address this issue, we have taken the following steps:

(1) Most literature-based statements previously included in the *Result Section* have been moved to the *Discussion Section*, where they are more appropriately contextualized. Here are the modified texts, now reading as: (Lines 382~394)

"Human activity datasets, generated by road maps, nightlight data, and forest cover change datasets, have been used to delineate intact and managed forests in the previous study. However, a notable ambiguity arises from the definition of managed forests, as these studies and datasets did not further differentiate between various types of forest management practices[49]. Despite this limitation, both the previous study[18] and our findings illustrate the damage of human management practices on vegetation resilience (Fig. 2). Managed forest fragments experience human disturbance and pressure, including forest planting, rotation, fragmentation, and irrigation. These factors directly influence various aspects such as tree species composition, forest age, and physiological processes[50], disrupting forest recovery processes. Furthermore, these management practices also affect how vegetation responds to climate change, ultimately diminishing the resilience of forest ecosystems in recovering from disturbances[18, 51]."

(2) For statements in the *Result Section*, we have ensured it is explicitly stated whether they are based on our findings or supported by the literature. Proper citations and clear connectors such as "consistent with" have been added for clarity. Here are the modified texts:

"The CSD theory indicates that systems with lower resilience are more sensitive to external disturbance[42]." (Lines 309~310)

"This analysis reveals a general decline in the resilience of global forests, which is consistent with previous studies[18, 48]." (Lines 348~349)

6.- The definition of forest types seems pretty central to the research question, so it should also be central in the text, but it is now only relegated to the supplementary materials. Either tables S1 and/or S2 should be moved to the main manuscript, or the most important forest type definitions should be clearly stated in the text.

R: We agree with your comment that these definitions are critical to the study and should be more prominently presented in the main text. **Action taken.** To address this, we have moved the Table S1 into the manuscript (Table 1), and briefly introduced the specific forest management types. Here are the briefly introduced the specific forest management types: (Lines 541~551)

"Agroforestry is a forest ecosystem that integrates trees with crops. Oil palms are tropical monoculture plantations for oil production, identifiable by their distinctive crown shapes. Shorter-rotation planted forests are established with fast-growing tree species for rapid biomass production, typically involving rotations of up to 15 years aimed at producing timber, pulp, or bioenergy. Longer-rotation planted forests are planted for rotations exceeding 15 years, focusing on ecological restoration, carbon sequestration, and biodiversity enhancement. Natural forests with human managements are subject to human interventions such as logging, trimming, and cutting, while natural forests without human disturbance are pristine ecosystems (e.g. Amazon Forest) that preserve their original biodiversity and ecological processes. The details for specific forest management types can be found in Table S1[1]."

Here are the introduced of Table 1: (Lines 562~571)

"We further re-classified forest types based on a classification of forest management types to compare resilience differences (Table 1). For instance, we compared resilience between natural forests (NF) and Plantations. Plantations are defined as all human-planted forests including planted forests (i.e., with longer rotating years and mainly for ecological restoration), short-rotation planted forests (up to 15 years), agroforestry, and oil palms. Reclassified forest types include NF WOM (natural forests without managements), NF WM (natural forests with managements), F WM (forests with managements), Plantations, PF and NF (Table 1). This classification allowed a detailed comparison of resilience across varying levels of human interventions."

Table 1. The definition and description of abbreviations for forest managements

Abbreviation	Definition	Description
NF	Natural forests	It concludes the natural forests without management and natural forests with management.
PF	Planted forests	It concludes the planted forests (longer rotating time over 15).
NF WOM	Natural forests without management/ Unmanaged natural forests	It concludes the natural forests without management (original or pristine forests)
NF WM	Natural forests with management/ Managed natural forests	It concludes the natural forests with management such as logging and clearing up.
F WM	Forests with	It concludes the planted

	management/ Managed forests	forests, planted forests with rotating up to 15 years, agroforestry, oil palms, and natural forests with management.
Plantations	Economic Plantations/ Board definition of planted forests	It concludes the planted forests, planted forests with rotating up to 15 years, agroforestry, and oil palms

7.- The arguments about the independence between forest management and resilience of the GFMD data should be more central in the discussion, as it is potentially a huge underlying limitation. Therefore, the independent tests on crowdsourcing data should be highlighted more.

R: We agree that this is an important aspect to involve in the *Discussion Section*. **Action taken.** We have added a new paragraph describing analysis based on crowdsourced data, and discussed the consistency and inconsistency with results by paired grids. Here are modified texts, now reading as: (Lines 469~491)

"The GFMD map generated by factors like LAI may affect the independence of our findings. This study extracts the resilience of crowdsourced points overlapped with the GHFC map. Results show that intensive human managements reduce resilience, aligning with our initial GFMD-based findings. For example, PF with rotations up to 15 years (i.e., wood plantations/shorter-rotation PF) exhibit lower resilience compared to those with rotation periods over 15 years (i.e., planted forests for restoration/longer-rotation PF) (Supplementary Fig. S26). These shorter-rotation PF are economically composed of a single species and are often partially harvested at regular intervals[66]. As a result, species richness and biodiversity in shorter-rotation PF are generally lower compared to longer-rotation PF, reducing the capacity to recover from disturbances[52]. Our study also supports the indications that the increase in tree species richness can further enhance forest resilience (Supplementary Fig. S26). It encourages multi-species plantations as a policy recommendation for the enhancement of forest resilience for afforestation. Additionally, the results by the nearest neighbor method showed that unmanaged NF demonstrate higher resilience than managed NF

(Supplementary Fig. S27), aligning with our initial GFMD-based findings. However, some findings differ from GFMD-based outputs. For example, the resilience difference between PF and NF remains below 0 until P/PET < 2.5 based on the crowdsourced datasets (Supplementary Fig. S28), while the GFMD-based findings in Figure 3 show this threshold occurs until P/PET < 1.5. This discrepancy is potentially due to the imbalance of samples in the crowdsourced points, where samples of forest managements like PF are less as compared to NF. Overall, caution is needed when interpreting specific thresholds and effects."

8.## Specific points

R: We sincerely appreciate your detailed comments and suggestions. We have addressed the relevant points one by one in our responses table and texts. To ensure clarity, we have categorized the responses into two parts: Part 1: the comments and suggestions about the figures, and Part 2: the comments and suggestions about the texts.

Part 1: the comments and suggestions about the figures

The comments and suggestion about the figures:

Fig 1: confusing and not vizually clear. strangely structured. Axis titles are missing. Lat/lon indicators are missing. So far, LAI has never been mentioned in the text, so the subplots b and d come as a bit of a surprise. Background grid lines could aid the comparison of subplots b, d, e., and the x-axes could be aligned for more consistency and ease of interpretation. In the figure description: "as an example of a specific global grid point", you could say what point you are presenting (e.g. central coordinates). Finally, the last part of the figure caption probably belongs to the main text: "The other examples in Central Africa… …found in Figure S2, showing similar spatial patterns".

R: We have corrected Figure 1 in the revised version of the manuscript according to the reviewer's suggestion. **Action taken.** Below are our responses and the corresponding changes made: (1) We have restructured the figure layout for better organization such as moving the color label to the bottom of the figure. (2) Background grid lines have also been added to subplots b, d, and e to aid comparison (add the Lat/Lon in the figures). (3) The x-axes have been aligned for more consistency for b, d, and e. We added the specific global grid point in the figure for explanation (-88.877, 31.905). (4) We added the introduction of LAI and resilience metrics (Lines 116~128 and Lines 129~135). We have revised the main text to introduce LAI earlier and explain its relevance to the analysis. We also introduce the resilience increases with the decrease of AR1/λAR1/λVar. (5) We agree that the last part of the figure caption belongs to the main text (Lines 139~142). This content has been relocated to the main text, and the figure caption has been revised according. Here are the modified texts:

"We derived the reference distribution of forest managements for the year 2000 (Fig. 1a) by overlapping tree cover data available for the year 2000 derived from the Global Hansen Forest Change Map (GHFC) and the spatial mapping of forest management types retrieved from the Global Forest Management Datasets (GFMD) for the year 2015. Time series of multiple vegetation indices were retrieved for the selected forest pixels over the period 2001-2015, including leaf area index (LAI), kernel normalized difference vegetation index (KNDVI), fraction of absorbed photosynthetically active radiation (FAPAR), gross primary productivity from MODIS (GPP_MODIS), and GPP from the Penman-Monteith-Leuning model (GPP_PML). Seasonal period patterns and long-term trends were removed from the original vegetation time series to make them stationary (Fig. 1b, d) and CSD indicators were retrieved on the residual signal to characterize the forest resilience. In this study, we express resilience in terms of AR1, λAR1, and λVar (additional details in the *Method Section*)." (Lines 116~128)

"Results show that areas characterized by human activity, as reflected in the forest management map, generally have lower resilience than natural forest without management (NF WOM) (Fig. 1). This emerges particularly evident for instance in the Amazon and Central African tropical forests where the AR1 estimated from LAI (AR1-LAI) in natural forests with management (NF WM) is substantially higher than that one quantified in NF WOM (Fig. 1j, k and Supplementary Fig. S1). The higher AR1/λAR1/λVar indicates lower forest resilience." (Lines 129~135)

"Resilience estimated by using other vegetation indices (KNDVI, FAPAR, GPP_MODIS, GPP_PML) and alternative CSD indicators (λAR1, λVar) exhibits similar spatial patterns (Supplementary Fig. S2) and confirm the robustness of our assessment." (Lines 139~142)

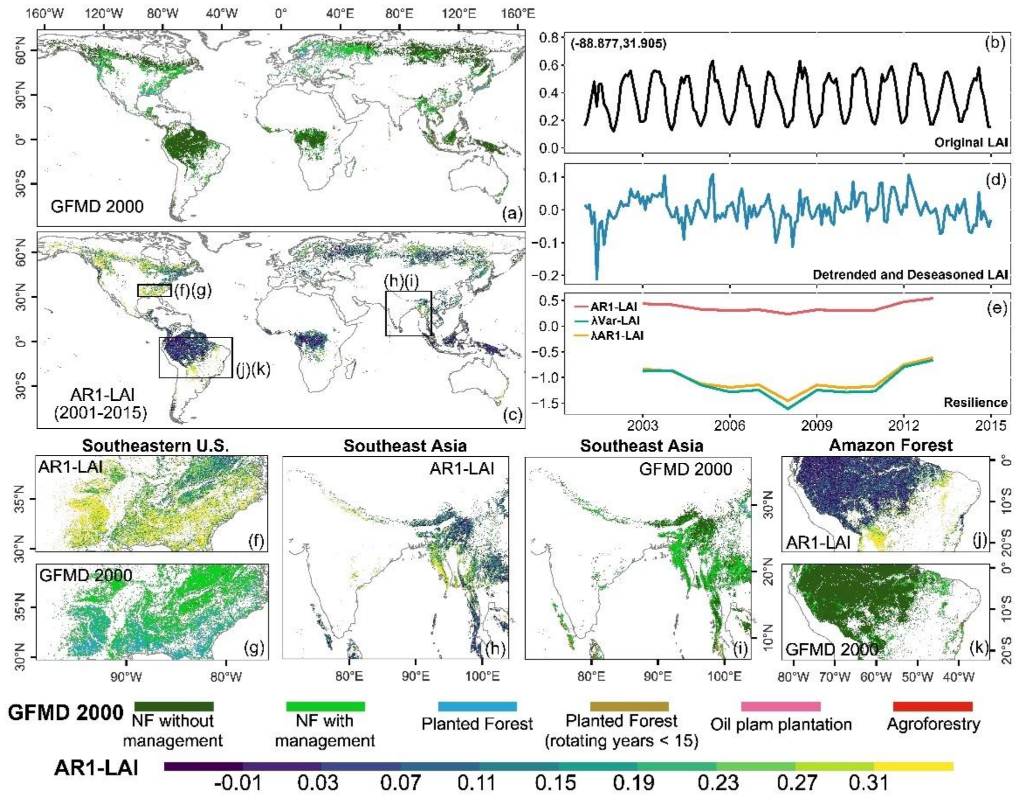


Fig. 1 Spatial Distribution of Global Forest Managements and Forest Resilience. (a) The spatial distribution of global forest management data (GFMD) in 2000 is generated by overlaying GFMD in 2015 onto GHFC in 2000 (details in the *Method Section*). (c) The spatial distribution of forest resilience displayed here is calculated using LAI and AR1 from 2001 to 2015 (AR1-LAI). (b-d) The original LAI time series is processed with detrending and deseasonalization methods to calculate resilience, using the specific global grid point (-88.877, 31.905) as an example. (e) Vegetation resilience in this study is estimated by AR1, λAR1, and λVar, showing similar changes at the specified grid point. The higher values of AR1, λAR1, and λVar indicate lower forest resilience. (f–k) Forest management activities are observed to reduce vegetation resilience in geographically similar regions, such as (f, g) the Southeastern United States, (h, i) South Asia, and (j, k) the Amazon Rainforest.

Figure 2: the color palettes chosen to represent the resilience difference and the datasets are very similar (browns and greens especially). This makes the interpretation of the figure difficult, so the colors should be differentiated more. Alternatively, subplots b, d, and f should come with their own color legend. In the figure description (here and in other figures) you mention a significance test,

but it is not specified what was tested. The significant difference from 0?

R: We have corrected Figure 2 in the revised version of the manuscript according to the reviewer's suggestion. **Action taken.** Below are our responses and the corresponding changes made: (1) We agree that the original color palettes for resilience differences and datasets were visually similar, making interpretation challenging. To address this, we have updated the color palettes (now the color palette is blue and red) in Figure 2. (2) In addition, we have added separate legends for b, d, and f subplots for reference. (3) Thank you for pointing out the need for clarification regarding the significance test. In the revised figure description, we have specified that the significance test evaluates whether the differences are significantly different from 0. Here are the modified texts and Figure 2:

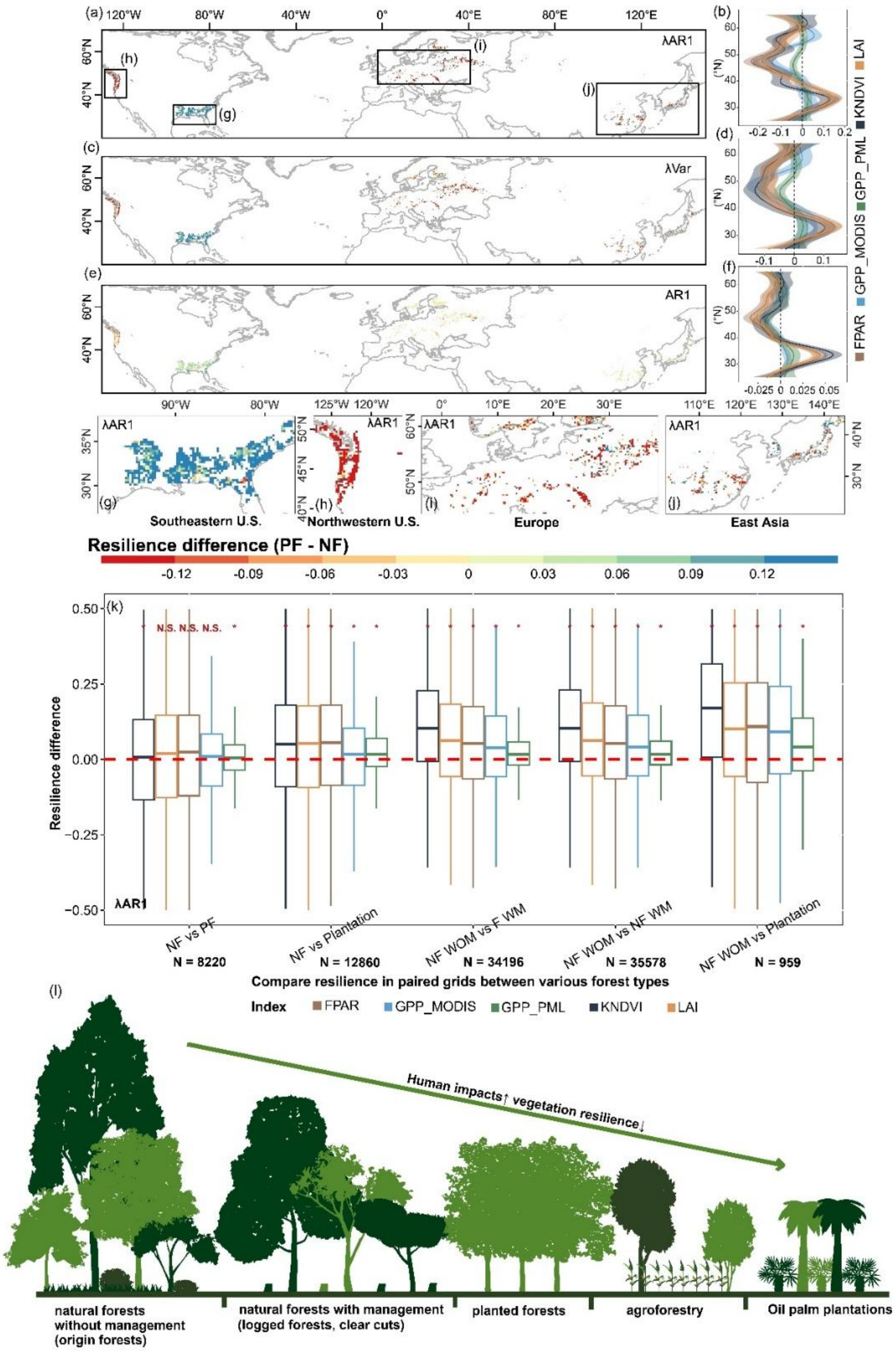


Fig. 2 Vegetation Resilience Differences Under Forest managements. (a-f) Spatial patterns of resilience difference between PF and NF are estimated using LAI and three measurements: (a, b) AR1, (c, d) λVar, and (e, f) λAR1. (b, d, f) These resilience differences are also assessed using multiple vegetation indices—GPP_MODIS (blue), GPP_PML (green), LAI (orange), KNDVI (blue-black), and FPAR (earthy color)—which display similar patterns across the latitudinal gradient. (g-j) Paired grids in key regions including the Southeastern U.S., Northwestern U.S.,

Europe, and East Asia are highlighted to showcase regional variations. (k) Further comparisons of resilience are performed across various forest management types using the same methodologies. These comparisons include NF vs. PF, NF vs. Plantations, NF WOM vs. F WM, NF WOM vs. NF WM, and NF WOM vs. Plantations. The notation "X vs. Y" indicates the resilience difference, calculated as the resilience of Y minus that of X. Statistical significance tests are performed to assess whether these resilience differences are statistically distinguishable from zero. Bars marked with an asterisk (*) denote statistical significance at $p < 0.05$, while bars labeled with 'NS' denote non-significant differences. (l) A conceptual figure illustrates different forest management types (e.g., NF WOM). Abbreviations used in the figure include: NF-Natural Forests, PF-Planted Forests, NF WOM-Natural Forests without Managements, NF WM-Natural Forests with Managements, and F WM-Forests with Managements.

Figure 4: I am a bit confused about SHAP values vs feature values. What do they represent and why do we need both? Are the feature values similar to effect coefficient? It should be made clearer even for readers that are not very familiar with the methodology.

R: Thank you for your question regarding SHAP values and feature values. (1) SHAP values represent the contribution of each feature to the model's prediction for a specific observation (i.e., the ranking of importance). (2) They also quantify how much a particular feature value increases or decreases the predicted resilience difference. Positive SHAP values indicate a positive contribution, while negative SHAP values indicate a negative contribution. Feature values are the actual input values (e.g., temperature, root depth) for a given feature in the dataset. They represent the observed or measured characteristics of the biophysical or forest management factors in the study. While SHAP values explain the contribution of features to the model prediction, feature values provide context for interpreting these contributions. For instance, higher temperature with SHAP values > 0 indicates its positive impact on resilience differences, but the feature value (e.g., a specific temperature range like higher temperature values) helps us understand the conditions under which this impact occurs. Thus, high temperatures have positive SHAP values and low temperatures have negative SHAP values, indicating that as the temperature increases, $\Delta\lambda$AR1 between PF and NF will also increase.

Action taken. We agree that providing a clearer explanation is essential for readers who may not be familiar with the methodology. Below, we clarify their roles in the *Result and Method Section*:

"The ranking of Shapley Additive Explanations (SHAP) values suggests that climate variables, such as temperature and P/PET, are more important than other factors like biomass and root depth (i.e., related to the vegetation context) (Fig. 4b). The SHAP value reflects the effects of variables on the predicted resilience difference (e.g., positive or negative). While the feature value interprets whether a feature is at a level (e.g., high temperature) that amplifies or diminishes resilience differences. For instance, negative SHAP values associated with high belowground biomass indicate that high belowground biomass shows negative effects on resilience differences (Fig. 4b)." (Lines 280~287 in *Result Section*)

"Specifically, SHAP values show how environmental variables (e.g., temperature and root depth) influence the predicted resilience difference. Positive values indicate the positive effects, and vice versa. Feature values, like observed temperatures, interpret the effects under specific conditions. The direction and magnitude of the variable effects depend not only on the SHAP value itself but also on the value of the feature. A high feature value with a positive SHAP value indicates that as the feature rises to a higher value, it contributes positively." (Lines 720~727 in *Method Section*)

We also modified the related descriptions of SHAP values in the figure's caption.

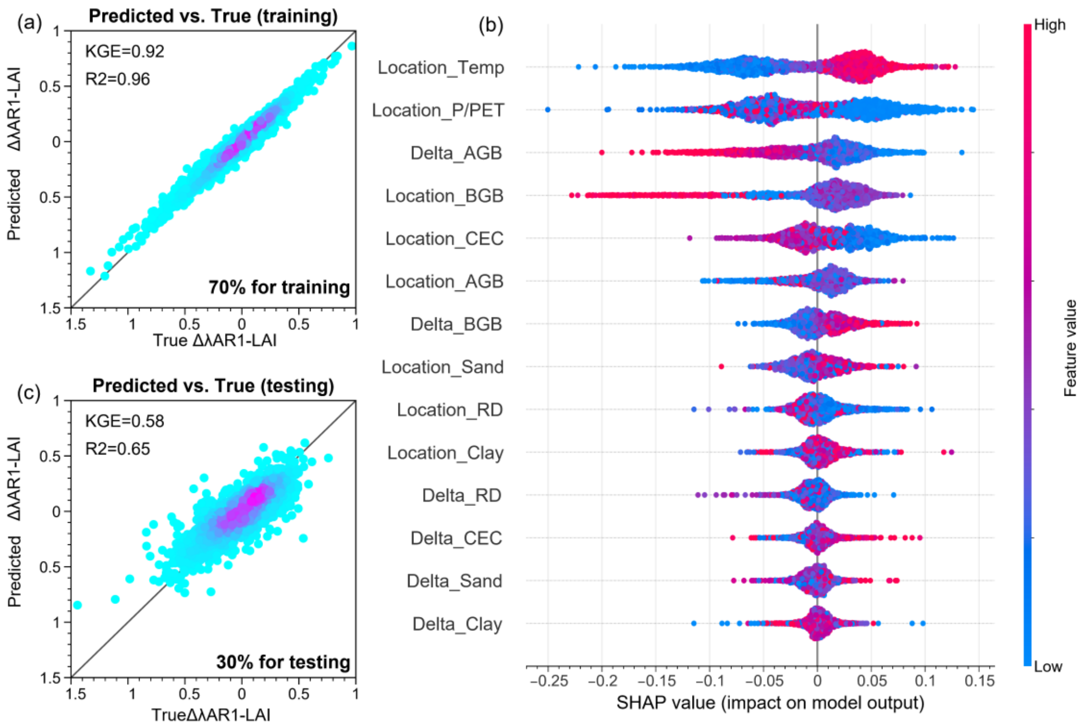


Fig. 4 Analyzing the Impact of Biophysical Factors on Vegetation Resilience Differences Between PF and NF. (a, c) Observed versus predicted $\Delta\lambda$AR1-LAI$_{(PF,\ NF)}$ are presented for both

training and testing samples. The coefficient of determination (R^2) and Kling-Gupta Efficiency (*KGE*) for each sample are indicated in the labels, with the frequency distribution depicted by the color gradient. (b) The SHAP values for environmental factors affecting the resilience difference are displayed. In this panel, red and blue points represent high and low feature values, respectively. SHAP values quantify the contribution of each feature to the model's prediction (e.g. positive or negative). For example, red points on the SHAP value plot that are greater than zero indicate the enhancing effects produced under high values. The "Location_Factor" indicates the local contexts of paired grid cells, while the "Delta_Factor" represents the difference in biophysical factors between PF and NF. Abbreviations of biophysical factors include: Temp-temperature, P/PET-the ratio of precipitation to potential evapotranspiration, AGB-above-ground biomass, BGB-below-ground biomass, CEC-cation exchange capacity, RD-root depth, Sand-sand fraction, and Clay-clay fraction.

Figure 5: Results descriptions and references to supplementary materials should probably go in the main text (results) and not in the figure description. Clarify what the significance test is testing. Also, "Vegetation greenness is limited by water in dry conditions and by energy in wet conditions, which may be related to shifts in resilience differences between PF and NF": to me, this sounds related to actual resource limitations rather than resilience differences.

R: Here, we respond to the reviewer's comment by addressing separately the two sections.

Section1: the comments on the caption of Figure 5

We have corrected Figure 5 in the revised version of the manuscript according to the reviewer's suggestion. **Action taken.** We agree that result descriptions and references to *Supplementary Materials* should be placed in the main text rather than in the figure caption (Lines 330~332). Furthermore, we have clarified the significance test in the caption, with the test of the significance between resilience differences and 0. We have updated the figure description to focus solely on the figure's content. Here are the modifications:

"The conceptual diagram illustrates the reversed resilience patterns of PF and NF, driven by vegetation sensitivity to climate factors (Supplementary Fig. S12)." (Lines 330~332)

"**Fig. 5 Climate Sensitivity of Vegetation for PF and NF.** (a, b) Spatial patterns of vegetation sensitivity to (a) SM and (b) PET for PF and NF are displayed. (c, d) Vegetation sensitivity is examined along a gradient of P/PET, emphasizing the variations in response across different climatic conditions. (e–j) Differences in climate sensitivity between PF and NF are quantified using five vegetation indices. Statistical tests are conducted to evaluate if these differences are statistically distinct from zero. Bars marked with an asterisk (*) denote significant differences ($p < 0.05$), while those labeled with "NS" indicate non-significant results. The climate variables analyzed include both water-related and energy-related factors: precipitation (Pre), soil moisture (SM), Palmer Drought Severity Index (PDSI), potential evapotranspiration (PET), and temperature (Temp). Standardization across these variables and vegetation indices ensures comparability of forest sensitivity to these factors." (Lines 334~344)

Section 2: the relationship between vegetation sensitivity and resilience.

Thanks for your insightful comment. We agree that the statement about vegetation greenness being limited by water in dry conditions and by energy in wet conditions does indeed reflect resource limitations, rather than directly addressing resilience differences between PF and NF. However, our intention was to highlight how these resource constraints might influence the resilience of the ecosystems. This difference in sensitivity to resource availability could help explain the observed resilience differences between PF and NF. (1) Vegetation sensitivity refers to the degree to which vegetation greenness responds to changes in external conditions, such as precipitation. For example, high sensitivity to precipitation implies that a reduction in rainfall or the occurrence of drought would lead to a significant decline in vegetation greenness. High sensitivity may indicate that vegetation greenness is more susceptible to external climate changes or sudden climatic events, resulting in greater fluctuations and lower stability (weaker resilience). For instance, PF are more sensitive to water factors like precipitation than natural forests. When external disturbances occur (such as drought events), the biomass reduction in PF is greater than in NF. Therefore, plantations require a longer time to recover to their initial state, resulting in lower resilience compared to NF.
(2) Additionally, existing studies have demonstrated that vegetation sensitivity can also be used to predict tipping points[10]. Sensitivity is considered a key functional trait of vegetation, and changes in sensitivity are often accompanied by shifts in resilience indicators, such as AR1. Critical Slowing Down (CSD) theory[11] indicates that as a system approaches a tipping point (TP), its sensitivity to external disturbances increases. This means that near the critical point, even small perturbations can

cause significant changes or large fluctuations in the system (i.e., Because the sensitivity of the system is very high, it's state can easily shift), a phenomenon known as "critical slowing down." In this context, the recovery time of the system is prolonged, indicating a decrease in the recovery rate. The reduction in recovery rate reflects the system's slowed pace in returning to its original stable state, making it more vulnerable to external changes.

Action taken. In response, we have modified texts in discussion to explain the relationship between sensitivity and resilience.

"Existing studies have indicated the relation between resilience and climate sensitivity[16, 60]. For instance, forests with higher sensitivity to water variability experience greater declines in biomass when subjected to external disturbances like droughts, prolonging the recovery process and making it more challenging for forests to recover their original state." (Lines 427~431)

L527: why are you limited to 1km by computational constraints? it seem very likely it could all be done at 500m

R: **Action taken.** Thanks for your important comments. In response, we tested the main conclusions at different resolutions such as 500m and 1000m. The results show that the resolution has a minor impact on our main conclusions. For example, the resilience estimates for parts of the African forest by 500m and 1km resolutions exhibit similar spatial patterns. The conclusions at 500m resolution also indicate that when P/PET < 1.5, the resilience of NF (Natural Forest) is greater than PF (Planted Forest); whereas when P/PET > 1.5, the resilience of PF is greater than NF. These findings are also aligned with the results by 1km. Overall, the spatial resolution of analysis datasets may not significantly influence our main findings (Fig .R2).

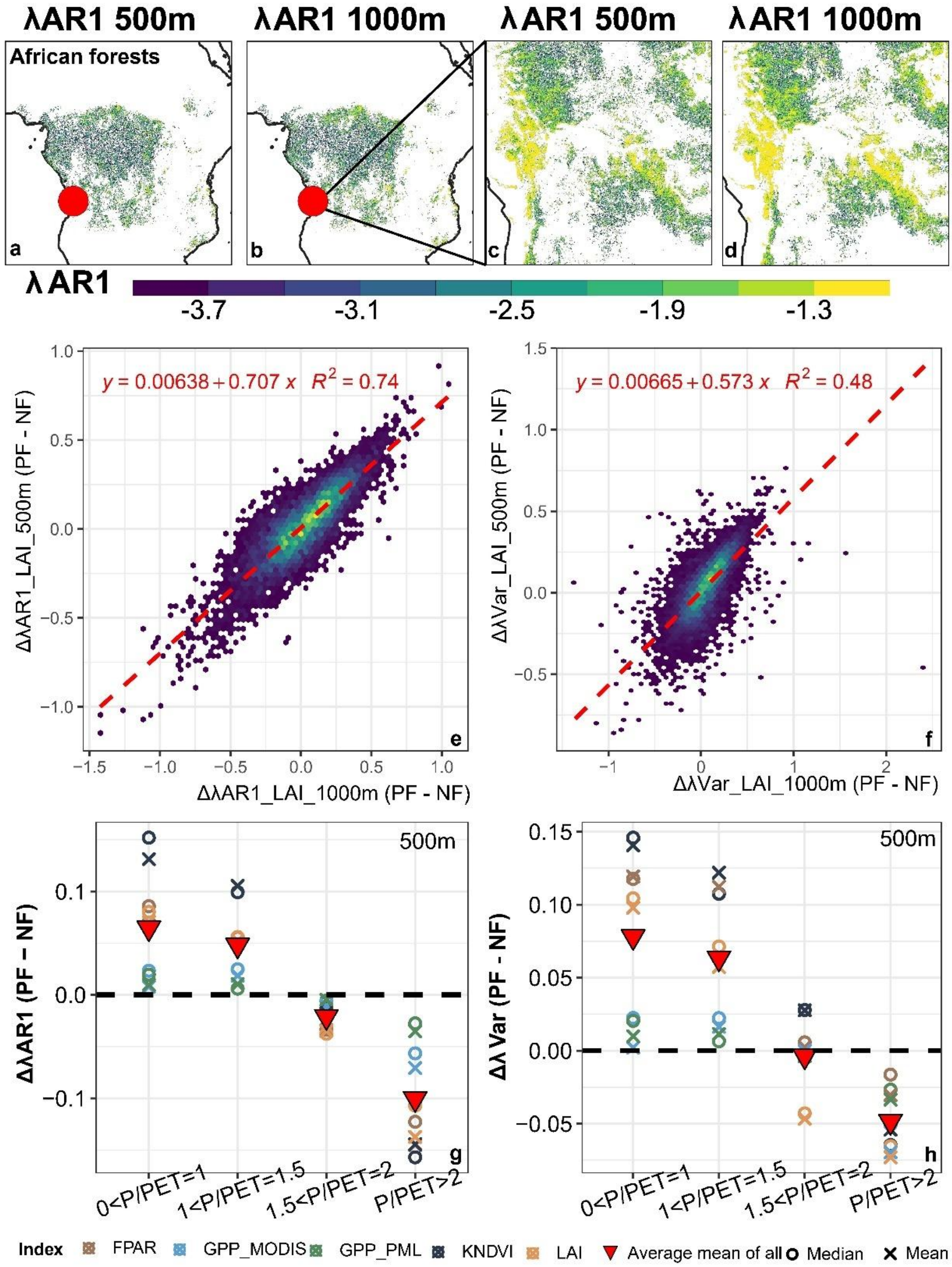


Fig R2. The results by 500m and 1000m spatial resolution. (a-d) The spatial distribution of λAR1 for the African forests by 500m and 1000m. (e, f) The resilience difference between PF and NF (λAR1 and λVar) by 500m and 1000m. (g, h) The resilience difference between PF and NF across the P/PET gradient.

Part 2: the comments and suggestions about the texts

L17: P/PEF should be defined

R: **Action taken.** We have added the definition of P/PET in the 15-17 lines of revised manuscript: "PF display higher resilience than NF in P/PET (the ratio of precipitation to potential evapotranspiration) > 1.5 (wetter climates), while NF is more resilient in P/PET < 1.5 (drier climates)."

L49: correct "bring" to "bringing".

R: **Action taken.** We rewrite this sentence. The updated sentence reads (Lines 48~50): "This is particularly critical in view of the ongoing intensification of natural and anthropogenic disturbance regimes that could affect the long-term resilience of global forests."

L60: "However, despite models appearing to predict stably, forest dynamics in observation remain complex and diverse" unclear sentence.

R: **Action taken.** We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The revised sentence now states (Lines 58~63): "Although these models are widely used for large-scale impact assessments, they only partially capture the complexity of forest dynamics, particularly the ecosystem's response to evolving human factors. Real-world forests may indeed undergo sudden and abrupt changes from a range of unexpected disturbances (e.g. forest clear-up by humans) that exceed model forecasts[27, 28]."

L62: correct "plating" to "planting".

R: **Action taken.** We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The updated sentence reads as (Lines 64~66): "activities like planting non-native species in arid zones not only resulted in unexpectedly lower tree survival rates but also led to a serious depletion of water resources[29]."

L64: remove double comma after "Furthermore"; clarify what is meant with "early stable predictions".

R: **Action taken.** We have addressed the issue by removing the double comma after "Furthermore" and clarifying the meaning of "early stable predictions" (i.e., the predictions of stable growth in Amazon forests from ecological models) for better readability. The updated sentence now reads as (Lines 66~67): "Furthermore, despite predictions of stable growth from ecological models, signs of death or degradation have been observed in the Amazon rainforest[30]."

L67: "Recent study report that…" correct singular/plural form (either "A recent study reports that…" or "Recent studies report that…".

L68: correct "dectect" to "detect"

R: **Action taken.** We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The updated sentence reads as (Lines 67~69): "A recent study reports that land surface models (i.e., Dynamic Global Vegetation Models) rarely detect the resilience trend in carbon sinks[31]."

L69-70: "…it will be better to…" this formulation sounds very strong. Both observational and modelling frameworks have their uses. I would present them as alternatives to one another, rather than stating that one is superior.

R: Thank you for your suggestion. We agree that both observational and modeling frameworks have their respective strengths and should be presented as complementary approaches rather than one being superior to the other. **Action taken.** The updated sentence now reads as (Lines 74~76): "Therefore, given the parameter and structure uncertainty of models, observational datasets offer an alternative to assess the impact of forest managements on the resilience of forests[35]."

L84-86: it might be worth spending a few more words/sentences to introduce and define the concepts of AR1, variance, λAR1, and λVar, as they seem rather central to the presented results.

R: **Action taken.** We agree that these resilience metrics need to be introduced. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The updated sentence now reads as (Lines 88~97): "The AR1 measures how strongly the current state of a system is influenced by its previous state[18], and the variance measures the magnitude of fluctuations around the system's equilibrium[43]. In addition, new metrics such as λAR1 and λVar have been developed, which are more closely related to the rate of vegetation recovery from disturbances, further aligning with the concept of resilience in dynamic systems[19, 44, 45]. Compared to the other three metrics, variance is susceptibly affected by environmental noise, and thus is challenging to detect critical transitions. As a result, recent studies have increasingly focused on AR1, λAR1, and λVar due to their stronger alignment with the concept of forest resilience[45, 46, 47]."

L90-93: "Motivated by the issue of what impacts of human activities especially forest management including PF and NF on vegetation resilience, this study used the latest global forest management dataset, multi-source satellite observations, and climate datasets to investigate. Here, this study mainly to:" unclear sentences.

L94: it might be worth introducing the "paired method" before mentioning it in the research questions.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We re-construct the sentence for clarification and introduce the paired grids method before the research questions. The revised version of the sentence is as follows (Lines 98~106): "In this study, we investigate how different forest management practices (e.g., PF, NF) interact with biophysical factors to affect global forest resilience. To achieve this, we integrate comprehensive global forest management datasets, multi-source satellite observations, and climate products. Forest resilience is expressed in terms of AR1, λAR1, and λVar from Critical Slowing Down (CSD) theory. To differentiate the impacts of human activities from natural variability, we employ a paired method that compares geographically proximate forest areas under different management practices. Machine-learning techniques are used to quantify the impacts of biophysical factors on resilience differences among forest management types."

L106: "In Spatial", meaning?

L107: "decline" as per meaning become smaller, fewer, weaker, etc. (https://www.oxfordlearnersdictionaries.com/definition/english/decline_1) is an intransitive verb, meaning that it cannot have an object. In your formulation: "…human forest management activities have declined the resilience" you want to use a transitive verb such as "decreased". Please check the use of decline vs decrease throughout the manuscript.

R: **Action taken**. The phrase "In Spatial" was unclear and unnecessary (in the original texts, we want to express “the results are based on spatial analysis”). We have removed it. We agree that "decline" is an intransitive verb and cannot take an object. Furthermore, we carefully reviewed the manuscript to change “decline” to “decrease”/”reduce” or change the “decline” to “…lower than…”. The revised sentence now reads (Lines 129~135):

“Results show that areas characterized by human activity, as reflected in the forest management map, generally have lower resilience than natural forest without management (NF WOM) (Fig. 1). This emerges particularly evident for instance in the Amazon and Central African tropical forests where the AR1 estimated from LAI (AR1-LAI) in natural forests with management (NF WM) is substantially higher than that one quantified in NF WOM (Fig. 1j, k and Supplementary Fig. S1). The higher AR1/λAR1/λVar indicates lower forest resilience.”

L108: Explain what the GHFC and GFMD are.

R: **Action taken**. We have added following explanations in the 116~119 lines: “We derived the reference distribution of forest managements for the year 2000 (Fig. 1a) by overlapping tree cover data available for the year 2000 derived from the Global Hansen Forest Change Map (GHFC) and the spatial mapping of forest management types retrieved from the Global Forest Management Datasets (GFMD) for the year 2015.”

L111: Rationale for these metrics is needed. We have no explanation of what they do, how they are constructed and what assumptions hide behind. This needs to be clear in the main text, not guessed or ignored. Also, in the introduction you talked about four warning signals: AR1, variance, λAR1, and λVar. Why are the results based only on three, excluding the variance metric?

R: The analysis excludes the variance metric for the following reasons: Variance is highly sensitive to external noise, such as climate variability and human activities, which can introduce bias. During critical transitions, variance can sometimes decrease instead of increasing, making it less reliable for detecting system state changes. Variance is more closely associated with system fluctuations, while AR1, λAR1, and λVar are more aligned with the dynamic concept of resilience and recovery rates in complex systems. Recently many studies have begun to focus on the application of λAR1 and λVar[4,7,8], due to their improvement from variance and AR1. **Action taken**. We have revised the manuscript to include detailed explanations and comparisons of AR1, variance, λAR1, and λVar in the *Introduction Section* (Lines 88~97).

L112: is the starting time of the analyzed data 2000 or 2001? This seems inconsistent.

R: The analysis starting from 2001 is based on the following reasons: (1) Sliding Window Analysis: MODIS Terra data in 2000 do not cover a complete calendar year, as observations began mid-year. This incomplete temporal coverage could affect the consistency of sliding window analysis and introduce biases in estimating seasonal variations, as different months would have unequal sample sizes. (2) Forest Cover Data Constraints: The Hansen Forest dataset provides forest cover data only for the year 2000. Estimation of forest cover has been applied more on Hansen's data and has a higher precision[12]. Considering these factors, we assume that the differences in long-term resilience estimation between 2001–2015 and 2000–2015 are negligible, and thus we chose to start the analysis from 2001 instead of changing the forest cover dataset.

L112-113: "The results revealed…" what results? It is not clear to me what analyses were performed. Are the statements in this paragraph purely based on visual inspection, or is there more to it that is not mentioned? This should be clarified.

R: **Action taken**. We rewrite these two paragraphs according to the reviewer's suggestion. The first paragraph now primarily describes the main methodology, clearly outlining the analytical framework and steps. The second paragraph focuses on discussing how forest management practices (e.g., plantation forestry, interventions in natural forests, and artificial forests) impact forest resilience in specific regions. Additionally, we have supported the conclusions based on visual inspection with relevant figures to enhance clarity and support the arguments. Here are the modifications (Lines 116~142):

"We derived the reference distribution of forest managements for the year 2000 (Fig. 1a) by overlapping tree cover data available for the year 2000 derived from the Global Hansen Forest Change Map (GHFC) and the spatial mapping of forest management types retrieved from the Global Forest Management Datasets (GFMD) for the year 2015. Time series of multiple vegetation indices were retrieved for the selected forest pixels over the period 2001-2015, including leaf area index (LAI), kernel normalized difference vegetation index (KNDVI), fraction of absorbed photosynthetically active radiation (FAPAR), gross primary productivity from MODIS

(GPP_MODIS), and GPP from the Penman-Monteith-Leuning model (GPP_PML). Seasonal period patterns and long-term trends were removed from the original vegetation time series to make them stationary (Fig. 1b, d) and CSD indicators were retrieved on the residual signal to characterize the forest resilience. In this study, we express resilience in terms of AR1, λAR1, and λVar (additional details in the *Method Section*).

Results show that areas characterized by human activity, as reflected in the forest management map, generally have lower resilience than natural forest without management (NF WOM) (Fig. 1). This emerges particularly evident for instance in the Amazon and Central African tropical forests where the AR1 estimated from LAI (AR1-LAI) in natural forests with management (NF WM) is substantially higher than that one quantified in NF WOM (Fig. 1j, k and Supplementary Fig. S1). The higher AR1/λAR1/λVar indicates lower forest resilience. Furthermore, lower resilience is also observed in Indonesian oil palm plantations and PF of southeastern United States compared to surrounding NF WOM (Fig. 1f, g and Supplementary Fig. S1). In Southeast Asia, NF WM and agroforestry demonstrate lower resilience relative to NF WOM (Fig. 1h, i). Resilience estimated by using other vegetation indices (KNDVI, FAPAR, GPP_MODIS, GPP_PML) and alternative CSD indicators (λAR1, λVar) exhibits similar spatial patterns (Supplementary Fig. S2) and confirm the robustness of our assessment."

L117-118: "Forest fragments experienced higher levels of human disturbances and pressures, such as logging, clearing, or agriculture 38" It should be clarified that these are results from the literature, if so.

L123-124: "These… influence…" check consistency of singular/plural forms; "management…" practices?

R: **Action taken**. We have revised the sentence to explicitly indicate that these findings are derived from previous studies. Then, we have moved the related literature discussion from the result to the discussion part. The revised sentences now read as (Lines 386~394):

"Despite this limitation, both the previous study[18] and our findings illustrate the damage of human management practices on vegetation resilience (Fig. 2). Managed forest fragments experience human disturbance and pressure, including forest planting, rotation, fragmentation, and irrigation. These factors directly influence various aspects such as tree species composition, forest age, and physiological processes[50], disrupting forest recovery processes. Furthermore, these management practices also affect how vegetation responds to climate change, ultimately diminishing the resilience of forest ecosystems in recovering from disturbances[18, 51]."

L125-127: Again, not entirely clear if these statements are based solely on studies from the literature or if they simply agree with the literature.

R: **Action taken**. We agree that it is essential to distinguish between conclusions drawn from our analysis and statements supported by the literature. To address this, we have revised the text as follows (Lines 131~134):

"This emerges particularly evident for instance in the Amazon and Central African tropical forests where the AR1 estimated from LAI (AR1-LAI) in natural forests with management (NF WM) is substantially higher than that one quantified in NF WOM (Fig. 1j, k and Supplementary Fig. S1)."

For results based on our analysis, we have now explicitly referenced the supporting figures and tables. For statements supported by the literature, we have included proper citations and used clear connectors such as " consistent with " to explain their context (Lines 348~349). We have moved the related literature discussion from the result to the discussion part (Lines 382~394).

L149: it is here we discover there are paired grids. But how are they paired? some info needs to percolate in the main text

L149: "paired grids" again, it should be briefly mentioned before what this method is about.

R: **Action taken**. According to your comments, we have revised the text to provide a clearer definition of the paired grids method in the *Introduction Section* (Lines 102~104). Additionally, we have specifically introduced the details of the method in *Result Section*. The revised texts in the *Result Section* now reads (Lines 159~165):

"Specifically, grids are paired using the sliding window method, which moves a fixed-size window across the study area. Within each window, paired grids are selected based on similar forest age (differences of ≤10 years), elevation (differences of ≤500 meters), and a minimum of 20% forest coverage for each management type (more details in the *Method Section*). We point out that the effects of potential covariates have been ruled out within the sliding window approach."

L168-169: “over 40% of paired grids show higher resilience in PF than in NF” this is quite high! If this is the case, I wonder if the statements about management decreasing resilience is founded or whether you should be more conservative. Considering the relevant uncertainties, is the 60%-40% ratio really different than a 50%-50% ratio based on random noise?

R: Thank you for this important question. Our intention is not only to highlight that certain management activities can diminish resilience, but also to emphasize the complexity and variability of these effects. While forest management activities associated with high human pressure—such as intensive plantations—can significantly reduce forest resilience, there are also management practices characterized by lower anthropogenic pressure that may enhance resilience under specific conditions. For instance, as illustrated in Figure 2, the most intensive form of human planting (i.e., plantations) shows the greatest negative impact on resilience, whereas less intensive afforestation practices (i.e., planted forests) have a relatively weaker effect on reducing resilience.

Action taken. Nonetheless, we recognize that the reader may be misled here, so we have rephrased the paragraph topic, now reading as (Lines 184~186). “Although forest management activities with strong human pressures have weakened forest resilience, managements with lower anthropogenic pressures can enhance resilience under specific conditions (Fig. 2).”

L174-175: “P/PET (the Ratio of Precipitation to Potential Evapotranspiration)” it is good that you define P/PET, but you should maybe also mention what it is since it is the first time you mention it in the text, i.e. an aridity index.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer’s suggestion (Line 16). In addition, we check similar issues throughout the manuscript.

L179: “these 15 resilience calculations…” not entirely clear what “these” refers to.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer’s suggestion. The modified version of the sentence now reads as (Lines 193~198): “To build confidence in these results, analyses have been produced for an ensemble of different combinations of CSD indicators and satellite vegetation indices (15 combinations = 5 vegetation indices × 3 CSD indicators). The overall high consistency emerging from the ensemble results confirms the robustness of the differences in resilience amongst the different forest management types investigated (Fig. 2a-f and Supplementary Fig. S4)."

L183: "Fig. S7", what happened to figure S6?

R: **Action taken**. We have ensured that all figures and supplementary figures are correctly referenced in the text and are presented in the proper order.

L221-223: "since… which…" correct the structure of the sentence. Currently the main clause is missing.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The modified version of the sentence now reads (Lines 231~233): "Notably, the number of paired grids in arid regions with P/PET less than 1 is insufficient (only 20% of grids are in the P/PET < 1 region), which may bring some uncertainties to the estimation of resilience."

L245: Mention what "background factors" are.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We change the "background factors" to "biophysical factors" based on the comment from you and Reviewer #4. We further explain it. The modified version of the sentence now reads as (Lines 265~267):

"The biophysical contexts including climate, vegetation, and soil variables may affect the resilience difference between PF and NF (details and models for other paired types in *Discussion*)."

In addition, we further described all related biophysics factors in the Figure 4 caption (Lines 304~307):

"Abbreviations of biophysical factors include: Temp-temperature, P/PET-the ratio of precipitation to potential evapotranspiration, AGB-above-ground biomass, BGB-below-ground biomass, CEC-cation exchange capacity, RD-root depth, Sand-sand fraction, and Clay-clay fraction."

L249: Correct "focued" with "focused".

R: **Action taken**. It has been corrected. The modified version of the sentence now reads (Lines 269~272): "Different vegetation indices and resilience estimation metrics show minimal impact on estimating the resilience difference (Supplementary Fig. S4); thus, we focus on the results for the λAR1 difference via LAI between PF and NF (ΔλAR1-LAI $_{(PF, NF)}$).”

L256: “explanatory” power?

R: **Action taken**. The explanatory power of model is demonstrated by its ability to predict outcomes accurately like R^2. We have corrected the text in the revised version of the manuscript according to the reviewer’s suggestion. The modified version of the sentence now reads (Line 27):

“with Model 1 having higher R^2 than Model 2 (Supplementary Fig. S9).”

L259: “other factors”, which ones? Maybe make a few examples.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer’s suggestion. The modified version of the sentence now reads (Lines 280~282):

“The ranking of Shapley Additive Explanations (SHAP) values suggests that climate variables, such as temperature and P/PET, are more important than other factors like biomass and root depth (i.e., related to the vegetation context) (Fig. 4b).”

L281-285: It is a bit strange that subplots e (PDSI) and i (temperature) are presented before a, b, c, etc.

L285: “…increases as the enhancement of…”, you mean “…increases with the increase of…”?

R: **Action taken**. We have adjusted the text for clarity and logical flow. Subplot references now follow a consistent order, and the “phrasing regarding the increase” has been corrected (Lines 310~315):

“We investigate the sensitivity of forests to water-related and energy-related factors such as soil moisture (SM) and PET (Fig. 5 and Supplementary Fig. S10). Results show that forest sensitivity to SM decreases with the increase of P/PET, while the sensitivity of forests to PET increases with the increase of P/PET (Fig. 5a-d). It indicates that forests in humid regions are limited by energy, while those in arid areas are limited by water resources.”

L287-288: “In this regard, results further suggests that PF are more sensitive to water resources than NF (Fig. 5e-5j)” this is a bit hard to follow, since the results are presented via sensitivity difference.

You could add a sentence to explain why this is the case.

R: **Action taken**. We have revised the text to explicitly explain why PF are more sensitive to water resources than NF, based on the sensitivity difference. The updated sentence now reads as (Lines 315~318):

"Furthermore, results reveal that the difference in vegetation sensitivity to water-related factors between PF and NF is over 0, suggesting that PF are more sensitive to water resources than NF in regions with P/PET < 1.5 (Fig. 5e-j)."

L291: what happened to Figure S10?

R: **Action taken**. We checked the full text and rearranged the figure's references in order. The revised sentences now read as (Lines 310~312):

"We investigate the sensitivity of forests to water-related and energy-related factors such as soil moisture (SM) and PET (Fig. 5 and Supplementary Fig. S10)."

L292: sensitivity to "soil water" or "soil moisture" might be more appropriate than "water" in general, as SM is what is shown in Fig. 5c).

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The modified version of the sentence now reads as (Lines 321~324):

"Although the sensitivity difference of forests to water-related factors between PF and NF is over 0 in regions with P/PET > 1.5 (Supplementary Fig. S11), forests sensitivity to water-related factors like SM in wetter climates is below 0 (Fig. 5c)."

L296-298: "Because the energy sensitivity difference between PF and NF gradually increases along with the wetter climate, and the variations of vegetation greenness is energy-limited in humid areas" missing a main clause.

R: **Action taken**. Based on our overall considerations, we have removed this sentence. We have provided a detailed description of the association between water sensitivity and resilience and then extended to energy factors. The modified version of the sentence now reads (Lines 321~327):

"Although the sensitivity difference of forests to water-related factors between PF and NF is over 0 in regions with P/PET > 1.5 (Supplementary Fig. S11), forests sensitivity to water-related factors like SM in wetter climates is below 0 (Fig. 5c). It suggests that water resources are not the primary factor influencing LAI in wetter climates. In contrast, NF shows a higher sensitivity to energy-related factors in climates where P/PET exceeds 1.5 (Fig. 5e-j), aligning with that PF may exhibit higher resilience than NF in these wetter climates (Fig. 3)."

L298-299: "As the enhancement of P/PET" does not seem grammatically correct. Change to "With increasing P/PET" or "As P/PET increases".

Action taken. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The modified version of the sentence now reads (Lines 327~332):

"Forests shift from being water-limited to energy-limited as P/PET increases, and this shift may explain the transition in resilience between PF and NF across varying climatic zones (Fig. 5 and Supplementary Fig. S10-12, details in *Discussion*). The conceptual diagram illustrates the reversed resilience patterns of PF and NF, driven by vegetation sensitivity to climate factors (Supplementary Fig. S12)."

L327: λPF – λNF is a newly introduced metric/abbreviation. Explain what it is. It is also not found in the referenced figure 5.

L332: "…PF and NF are reducing in both…" should be "decreasing" instead of "reducing".

R: **Action taken**. We have changed the λPF – λNF to $\Delta\lambda$AR1-LAI$_{(PF,\ NF)}$. $\Delta\lambda$AR1-LAI$_{(PF,\ NF)}$ indicates the resilience difference between PF and NF, that is estimated by λAR1 and LAI index. We also update these using in the full texts (like Lines 167~176). We have modified the "reducing" to "decreasing". We have addressed the following issues, now reading as (Lines 353~360):

"Specifically, the $\Delta\lambda$AR1-LAI $_{(PF,\ NF)}$ decreases in drier regions (P/PET < 1.5) with a trend of -0.003 yr^{-1}, while it increases in wetter areas (P/PET > 1.5) with a trend of +0.002 yr^{-1} (Fig. 6c-e). Considering the direction of resilience differences between PF and NF ($\Delta\lambda$AR1-LAI $_{(PF,\ NF)}$ >0 in P/PET < 1.5, and $\Delta\lambda$AR1-LAI $_{(PF,\ NF)}$ <0 in P/PET > 1.5), these findings indicate that the absolute value of $\Delta\lambda$AR1-LAI decreases in both drier and wetter zones from 2003 to 2013 (Fig. 5). This implies that the difference between PF and NF reduces over time."

L359: "However, a notable ambiguity arises from the definition of managed forests also came to this study" unclear sentence.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion (we removed the literature discussion in the *Discussion Section*). The modified version of the sentence now reads (Lines 382~386):

"Human activity datasets, generated by road maps, nightlight data, and forest cover change datasets, have been used to delineate intact and managed forests in the previous study. However, a notable ambiguity arises from the definition of managed forests, as these studies and datasets did not further differentiate between various types of forest management practices[49]."

L388: "CEC" is not defined so far.

R: **Action taken**. According to the reviewer's suggestion, we have described the effects of CEC on resilience difference between PF and NF in the *Result Section*, now reading as (Lines 288~291):

"Results of SHAP values indicate that a wetter and colder climate, higher levels of aboveground and belowground biomass, and a higher soil cation exchange capacity (reflecting higher soil fertility) may minimize the resilience differences between PF and NF (Fig. 4b)."

In addition, we have added the definition of the abbreviation in the 304~307 lines of the Figure 4 caption.

"Abbreviations of biophysical factors include: Temp-temperature, P/PET-the ratio of precipitation to potential evapotranspiration, AGB-above-ground biomass, BGB-below-ground biomass, CEC-cation exchange capacity, RD-root depth, Sand-sand fraction, and Clay-clay fraction."

L399-401: repeated sentence.

R: **Action taken**. We have deleted the repeated sentence.

L416-418: unclear sentence.

R: **Action taken**. We have rewritten the text in the revised version of the manuscript according to the reviewer's suggestion. The modified version of the sentence now reads (Lines 436~441): "Trees in PF have minimal height variation, which enables them to receive similar levels of solar radiation and create a stable microclimate to counter temperature changes[63]. It allows PF to maintain stable physiological processes under fluctuating energy conditions. NF exhibit significant height difference, with shorter trees often unable to access adequate radiation due to shading by taller trees[64]."

L420-422: "More interestingly, vegetation greenness is water-limited in dry regions (PF is more sensitive to water) and energy-limited in wet regions (NF is more sensitive to energy)". To me, this sounds rather obvious. It is well known and intuitive that vegetation is limited by water where water is scarce, and energy becomes the limiting resource where water is abundant. Maybe you want to reformulate this to frame it more in the context of resilience and sensitivity.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We have added an example to explain the relationship between sensitivity and resilience. The modified version of the sentence now reads as (Lines 427~431):

"Existing studies have indicated the relation between resilience and climate sensitivity[16, 60]. For instance, forests with higher sensitivity to water variability experience greater declines in biomass when subjected to external disturbances like droughts, prolonging the recovery process and making it more challenging for forests to recover their original state."

L438: I believe one needs to reach this point to learn for the first time that these indices are made from LAI

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. To address this, we have revised the *Result Section* to better introduce these metrics and vegetation products (Lines 116~128).

L451: "decline" change to "decrease".

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. Since decline is an intransitive verb, we have checked the whole manuscript and replaced "decline" with "decrease" or "reduce".

L458: "the primary focus is on the differential impacts of forest management between PF on resilience" unclear sentence.

R: **Action taken**. We have rewritten the text in the revised version of the manuscript according to the reviewer's suggestion, now reading as (Lines 514~520): "This study prominently examines the differences in resilience between PF and NF, further focusing on their response to biophysical factors. Certain factors (e.g., landforms and slope), that influence the vegetation recovery process, may not have been adequately accounted for in the model[70]. Furthermore, we also built XGBT models to investigate the influence of biophysical contexts on the resilience differences between NF and Plantations, NF WM and NF WOM, NF WOM and F WM, and between NF WOM and Plantations (Supplementary Fig. S29)."

L480: "…and reduce uncertainties from land use change…" and following sentence: to me, this sounds more like land use change was ignored since you are only assuming that pixels do not change, but in reality, they might be changing. Am I missing something?

R: Thank you for your important question. Vegetation changes can influence the comparison of resilience under different forest management practices. Therefore, it is necessary to minimize this influence. We achieved this by overlaying the GFMD and GFHC maps. If a pixel was classified as a forest in both 2000 and 2015, we assume/consider that no land-use change occurred in the pixel. Then, we only analyze pixels with no landcover change. **<u>Action taken</u>**. We have corrected the text, now reading as (Lines 558~560): "This process ensured the spatial consistency of the pixels during 2000-2015, and pixels with land cover change may not be accounted for in the analysis."

L491: "ASCE" define/explain what it means.

R: **<u>Action taken</u>**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion, now reading as (Lines 576~580): "we used monthly mean precipitation, maximum and minimum air temperature at 2 m, soil moisture, PDSI (Palmer Drought Severity Index), and potential evapotranspiration (PET that based on the American Society of Civil Engineers Penman-Monteith equation[75]), acquired from the TerraClimate database at 4-km (1/24th degree) for 2001-2015 period[76]. "

L503: change "photosynthetic" to "photosynthesis".

R: **<u>Action taken</u>**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion, now reading as (Lines 590~591): "Different vegetation indices were used to capture the forest biomass, greenness, and photosynthesis."

L506: "PML" and "MODIS" you may want to explain the abbreviations or mention the source (which for MODIS you do a few lines below).

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion, now reading as (Lines 591~596): "Here we selected five vegetation indices including LAI (leaf area index), FPAR (fraction of photosynthetically active radiation), GPP_PML (gross primary productivity from Penman-Monteith-Leuning model), GPP_MODIS (gross primary productivity from Moderate Resolution Imaging Spectroradiometer), and KNDVI (kernel-based normalized difference vegetation index) to monitor the dynamic changes in vegetation from 2001-2015. "

L511: this is a wrong use of kNDVI. see https://doi.org/10.1016/j.isprsjprs.2022.12.019

R: There are various methods for calculating KNDVI. The formula for calculating the KNDVI is: (KNDVI = tanh ((NIR - Red)/2σ)^2) [13]. Where NIR is the near-infrared reflectance, Red is the red band reflectance, and σ is a scaling parameter that normalizes the input. Previous studies have demonstrated that σ = 0.5(NIR + red) is a reasonable choice for related calculations[14]. Then, the equation can be simplified as KNDVI = tanh($NDVI^2$). This formulation ensures that the scaling parameter σ adapts dynamically to the reflectance values of NIR and Red, thereby preserving the physical interpretation of the vegetation index while minimizing sensitivity to noise in the reflectance data. The inclusion of σ accounts for variability in reflectance magnitudes, making the index robust to differences in illumination and sensor calibration. In addition, many previous studies used the simplified equation of KNDVI[4,15]. **Action taken**. We have corrected the text in the revised version, now reading as (Lines 597~598): "It is calculated using the simplified formula KNDVI = tanh($NDVI^2$),"

L516: "which are mainly from [MODIS]" and where are the other products from?

R: **Action taken**. We rewrite the description of the vegetation products according to the reviewer's suggestion, now reading as (Lines 599~608):

"These vegetation indices were derived from various vegetation products. Specifically, GPP_PML is derived from the PML_V2.0.1.7 model[82], and GPP_MODIS is derived from MODIS Terra (MOD17A2H.061) and Aqua (MYD17A2H.061) observations. These datasets offer a spatial resolution of 500 meters and an 8-day temporal resolution. LAI and FPAR are obtained from MYD15A2H.061 and MOD15A2H.061 MODIS Aqua and Terra observations, also at a 500 m spatial resolution and 8-day temporal resolution. NDVI is derived from MOD13A2.061 and MYD13A2.061 MODIS Terra and Aqua observations, with a 1 km spatial resolution and 16-day temporal resolution. The details of these products can be found in *Supplementary Materials* (Table S2)."

L518-519: "these two observed vegetation greenness or biomass are similar in a day" give a reference stating this.

L519: "But two times may be affected by the cloud" unclear sentence.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion, now reading as (Lines 608~611):

"Furthermore, the Terra and Aqua satellites observe the earth's surface at 10:30 am and 1:30 pm, and their observations of vegetation greenness or biomass are similar on a given day[83]. These datasets may be affected by clouds, snow, and shadowing."

L524-526: You mention two contrasting spatial resolutions for the vegetation indices (500 m and 1 km). Which one is it?

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We have revised the text to ensure clarity and explicitly state the temporal and spatial resolutions for all datasets (Lines 599~608).

L530: you still have not explained how to calculate resilience. why not start with that before?

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We have placed the Estimating Vegetation Resilience subsection after the dataset's subsection, followed by matching methods (Lines 617~706).

L531: If not already mentioned in the introduction, the "paired grids" concept should really be explained in detail here, as I am still wondering what it means.

Action taken. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We agree that this is a critical methodological element and have ensured it is explained in detail across the manuscript, specifically in the *Introduction, Result, and Method Sections*.

In the Introduction:

We introduced the paired grids concept as part of the methodology in the *Introduction Section* (Lines 102~104):

"To differentiate the impacts of human activities from natural variability, we employ a paired method that compares geographically proximate forest areas under different management practices."

In the *Result Section*:

We elaborated on how paired grids were constructed, specifying the use of the sliding window approach to ensure similarity in environmental conditions (Lines 159~165):

"Specifically, grids are paired using the sliding window method, which moves a fixed-size window across the study area. Within each window, paired grids are selected based on similar forest age (differences of ≤10 years), elevation (differences of ≤500 meters), and a minimum of 20% forest coverage for each management type (more details in the *Method Section*). We point out that the effects of potential covariates have been ruled out within the sliding window approach."

In the *Method Section*:

We have added the definition and significance of paired grids. In the *Method Section*, we have introduced the details of the paired grid method (Lines 679~706).

L538: what is Simon et.al?

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The modified texts now read as (Lines 690~691): "The global forest age data are obtained from Besnard et al.[88], while elevation data were derived from the Shuttle Radar Topography Mission (SRTM) v4 dataset[89]."

L541-543: mention supplementary figures explicitly.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. In response, we have reviewed the manuscript thoroughly to ensure all references to supplementary figures are explicitly stated and correctly cited.

L546: why monthly data? to do a robust time series analysis it is better to take all data available, not picking a single MVC value per month

R: We agree that robust time series analysis benefits from utilizing as much temporal data as possible, and we appreciate your suggestion. To address this concern, we conducted a sensitivity analysis to explore the impact of different temporal preprocessing methods on resilience estimation (Lines 765~781 for the methods). We tested the impact of the three preprocessing methods on the results, including (1) obtaining monthly maximum vegetation data via MVC from MODIS Terra and Aqua satellites, (2) acquiring original temporal resolution vegetation data from Terra satellite, (3) and obtaining original temporal resolution vegetation data from Terra satellite only for the growing season. Our results indicate that the choice of temporal preprocessing method does not significantly affect the main conclusions of our study. The related sensitivity test can be found in Supplementary Fig. S20-S23, and the related discussion can be found in Lines 457~462 and *Supplementary Note 2*.

Action taken. We have corrected the text in the revised version to mention the sensitivity test (Lines 674~677).

"Some CSD studies even suggest that missing values should be excluded in resilience estimating, as this could induce new uncertainty by interpolation methods[45]. In this regard, we conducted additional tests to evaluate the impact of data pre-processing methods on our key findings (Supplementary Fig. S20-S23)."

L550: you need to explain what these are

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion, now reading as (Lines 636~639): "Consequently, we evaluated forest resilience using AR1, λAR1, and λVar. Notably, these λ metrics (λAR1 and λVar) capture resilience within a linearized framework of the system's dynamics near a stable equilibrium, which can be formally modeled as an Ornstein–Uhlenbeck process."

L552: It is still unclear to me what the difference between AR1 and λAR1 is.

R: In fact, AR1 and λAR1 are closely related. Mathematically, λAR1 is simply the logarithm of AR1. In other words, λAR1 excludes cases where AR1 is less than 0. In ecological systems, recovery rate indicators typically assume that the system exhibits positive autocorrelation, meaning that past data influence current or future conditions (for example, if a drought leads to a decline in greenness in an earlier period, the drought's effect on greenness can still be observed in later data). However, in the dynamic real world, conditions can be complex and variable. Therefore, we retain both AR1 and λAR1 for analysis (because we also need to test this assumption). Here are the original texts to describe the relation between AR1 and λAR1 (Line 665): "λAR1 and λVar may result in null values due to logarithmic constraints (i.e., AR1 < 0 or $\frac{\emptyset^2}{Var[X]}$>1)."

Formulas: try to keep the font of parameters consistent between the formulas and the main text.

R: **Action taken**. We have addressed this issue by reformatting all formulas to ensure consistency. The font for parameters in the formulas has been standardized to Cambria Math, with a font size equivalent to "12 pt". We have reviewed the entire manuscript to ensure that the parameters in the formulas match those in the main text, both in terms of notation and style (see equations).

L571: "same as the Xn" unclear.

R: **Action taken**. We checked the full text for italicized notations about formulas and variables. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion, now reading as (Lines 650~651): "Where X_{n+1} denotes the system state at discrete time step $n+1$ (where X_n is the system state at discrete time step n)"

L591-593: correct typing mistakes: "however", "maxium", "northern".

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion.

L591: Explain what "MVC" stands for.

L592: Explain what "month datasets" are, or reformulate.

L593: "influencing" reformulate, unclear.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We have addressed the following points: MVC stands for Maximum Value Composite (we first introduce the MVC in Line 614), a method commonly used to handle missing data in satellite-derived vegetation indices. This has been clarified in the revised text. The phrase "month datasets" has been reformulated for clarity as "monthly vegetation datasets." The term "influencing" has been replaced with "affect," and the sentence structure has been adjusted to enhance readability and clarity, now reading as (Lines 671~672):

"Furthermore, to mitigate the impact of missing values on resilience estimation, we applied the MVC method to generate monthly vegetation data."

L595: "some CSD studies did not deal with missing values", does that mean that missing values are excluded or kept as they are? Either way, it is not possible to "not deal with missing values".

R: **Action taken**. We agree that the phrase "not deal with missing values" is unclear and could lead to misunderstanding. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The revised sentence now reads (Lines 674~677):

"Some CSD studies even suggest that missing values should be excluded in resilience estimating, as this could induce new uncertainty by interpolation methods[45]. In this regard, we conducted additional tests to evaluate the impact of data pre-processing methods on our key findings (Supplementary Fig. S20-S23)."

L603-605: "Not only because of the strong relationship between resilience difference between PF and NF and background variables, …" missing main clause. Also, this part specifically sounds as if you chose the method based on the results because "it looks good", instead of basing the decision on hypotheses or research gaps.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. Here are the modified texts, now reading as (Lines 708~713):

"We first explored the biophysical drivers for resilience differences between PF and NF. This is particularly important given the widespread attention to restoration activities[21, 22, 23], such as the UN Decade on Ecosystem Restoration (2021–2030). Understanding the resilience differences between PF and NF can provide valuable guidance for designing and implementing effective restoration strategies to enhance ecosystem resilience."

Additionally, following the suggestions of Reviewer #4, we have included comparisons involving other forest management types. Therefore, the main text focuses primarily on the differences between PF and NF, while comparisons of other types are included in the *Discussion Section* (Lines 514~533) and *Supplementary Materials (Supplementary Note 4)* with relevant descriptions.

L607: "widely utilized in ecosystem studies", then maybe give a few more references instead of only one.

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. Here are the modified texts, now reading as (Lines 715~717):

"XGBT combines gradient learning and tree-boosting steps, making it a powerful ensemble learning algorithm widely utilized in ecosystem studies[39, 90, 91]. "

L612: what does "its" refer to?

L615" what does "their" refer to?

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. Here are the modified texts, now reading as (Lines 728~733):

"We used XGBT models to predict the λAR1 difference between PF and NF (e.g., $\triangle\lambda\text{AR1-LAI}_{(\text{PF,NF})} = \lambda\text{AR1-LAI}_{(\text{PF})} - \lambda\text{AR1-LAI}_{(\text{NF})}$), based on eight environmental factors and the differences in these factors between PF and NF. The AR1 difference between PF and NF was the target variable, while the eight environmental factors (e.g., CEC value of the grid) and their differences (e.g., $\triangle\text{CEC}_{(\text{PF,NF})} = \text{CEC}_{(\text{PF})} - \text{CEC}_{(\text{NF})}$) between PF and NF were used as predictors."

L621: "datasets availability section" should be corrected to "Data and code availability section".

R: **Action taken**. Based on the feedback from yours and Reviewer #3, we have added descriptions of the relevant variables in the *Method Section*. The revised sentence now reads (Lines 735~740):

"The eight environmental factors include annual temperature (Temp derived from TeerraClimate), dryness index (P/PET derived from TeerraClimate), root depth (RD[92]), above-ground biomass (AGB[93]), below-ground biomass (BGB[93]), clay fraction (Clay[94]), sand fraction (Sand[94]), and cation exchange capacity (CEC[94]). These factors are selected for their known influence on forest resilience, including both climatic and land-surface variables. ”

L671: “low probability pixels” what does this refer to, low probability of what?

R: **Action taken**. To clarify, this refers to pixels with a low probability of accurate classification within the Global Forest Management Dataset (GFMD). Specifically, these are pixels where the classification accuracy of forest management types (e.g., planted forests, natural forests) is below a certain threshold (50%, 60%, 70%) based on the GFMD’s internal validation metrics. We have corrected the text in the revised version of the manuscript according to the reviewer’s suggestion, now reading as (Lines 788~791): “By removing low-accuracy pixels (i.e., low accuracy of classification in GFMD), the results of reanalysis displayed a similar pattern, and different accuracy thresholds may have a limited impact on our key conclusions.”

L681-683: “GFMD is the result of a classification algorithm that takes in input various factors including vegetation data, the same data to calculate resilience. It may undermine the independence between forest management and resilience”. This sounds like a potentially huge issue. If the input data are not independent, how can you say anything about the results? The following parts were you control with independent data should really be mentioned more centrally.

L702-703: “This is aligned with our main conclusion. The above analysis results can be seen in the discussion and supplementary section” good, as mentioned above, I think this confirmation should be more central to the results. Also, I could not find it in the discussion, where it should really be, rather than hidden at the end of the methods or in the supplementary materials.

R: **Action taken**. Thank you for your valuable suggestions. We agree that the discussion on the independence and reliability of the GFMD data should be more critical to the *Discussion Sections*. The specific modifications, including details of the expanded discussion, are outlined in the above-mentioned response (Modified texts can be read in Lines 496~491 of the *Discussion Section*).

L707: the data here means the data produced by the paper, not just the data used as input (although the latter is also needed)

R: **Action taken**. We updated the core datasets for the figure’s reproduction. We have corrected the text in the revised version of the manuscript according to the reviewer’s suggestion. Here are the modified texts: (Lines 835~838)

“The core part of datasets related to the figures and maps are now can be found: https://figshare.com/articles/dataset/_b_Biophysical_factors_determining_the_effects_of_forest_management_on_ecosystem_resilience_b_/28286651”

Response to Reviewer #2

Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

R: We greatly appreciate the time and effort you have dedicated to reviewing our manuscript and providing important feedback. We have modified and responded point-by-point accordingly.

Response to Reviewer #3

Reviewer #3 (Remarks to the Author):

1. This study explores the resilience of forests under varying management intensities, ranging from unmanaged natural forests to managed, planted forests and plantations. The research examines how resilience, as defined by the critical slowing down (CSD) theory, differs among forests with different levels of management. The study concludes that natural, unmanaged forests exhibit greater resilience, while planted forests are more vulnerable and thus show lower resilience. Additionally, the study investigates the factors influencing these differences, reporting that wetter regions are less vulnerable and, consequently, more suitable for planted forests.

I find the study both interesting and novel, contributing important findings beyond previous similar studies. However, several points need to be addressed to strengthen the results. The authors conducted multiple sensitivity analyses to demonstrate that the findings are consistent, considering factors such as the sliding window width, forest coverage threshold, definitions of planted (PF) and natural forests (NF), reliability of forest management maps, and preprocessing methods of satellite vegetation data. However, one crucial aspect that should be considered is forest age: natural and managed forests are typically older than planted forests and plantations, and resilience is known to vary among different age groups. Therefore, I recommend including an additional sensitivity analysis to compare planted forests with younger unmanaged stands, to verify that the results are not merely due to age differences between PF and NF.

R: We sincerely appreciate your detailed comments and constructive feedback. Below, we address your key suggestions on the forest age. We acknowledge that forest age is an important factor influencing resilience, and the differences between younger and older forests could play a role in the resilience comparison between planted forests (PF) and natural forests (NF). In this study, we have already accounted for the potential impact of forest age on our findings (Lines 687~691). In the *Method Section*, we used the paired grids method to compare vegetation resilience differences among different forest management types. We have excluded the paired grid with more than 10 years of forest age difference. For example, if the age difference between a paired PF and NF exceeded this threshold, the grid was excluded from the analysis. Here are the original texts in the main text: "To mitigate the potential effects of varying forest age and elevation on vegetation resilience, we ensured that the differences in these variables within each paired grid were restricted to within 10 years and 500 meters, respectively[16]. The global forest age data are obtained from Besnard et al.[17], while elevation data are derived from the Shuttle Radar Topography Mission (SRTM) v4 dataset[18]."

Action taken. We noticed that the critical information about paired grids and forest age difference was mentioned only once in the *Method Section*, and the manuscript did not adequately introduce or emphasize this important methodological approach. To ensure readers do not overlook these key details, we have mentioned the description of the paired grids method and the age thresholds in the *Result Section*. These revisions include as following: (Lines 159~165)

"Specifically, grids are paired using the sliding window method, which moves a fixed-size window across the study area. Within each window, paired grids are selected based on similar forest age (differences of ≤10 years), elevation (differences of ≤500 meters), and a minimum of 20% forest coverage for each management type (more details in the *Method Section*). We point out that the effects of potential covariates have been ruled out within the sliding window approach."

Additionally, according to the reviewer's comment, we conducted a sensitivity analysis by applying different age difference thresholds (e.g., 5, 7, and 10 years) to filter paired grids. The results consistently showed similar findings: forest management tends to weaken resilience, but under certain conditions, well-managed PF exhibited resilience levels comparable to or even exceeding those of NF (Supplementary Fig. S19). Importantly, the findings remained robust regardless of the age threshold applied. Here are modified texts and figures: (Lines 753~757)

"Specifically, we tested the impact of the sliding window widths (0.25°, 0.35°, and 0.45°), the threshold of forest age differences (5, 7, and10 years), and forest coverage thresholds (10%, 15%, and 20%) on the results, and found that they did not affect the key conclusions (Supplementary Figs.

S17-S19)."

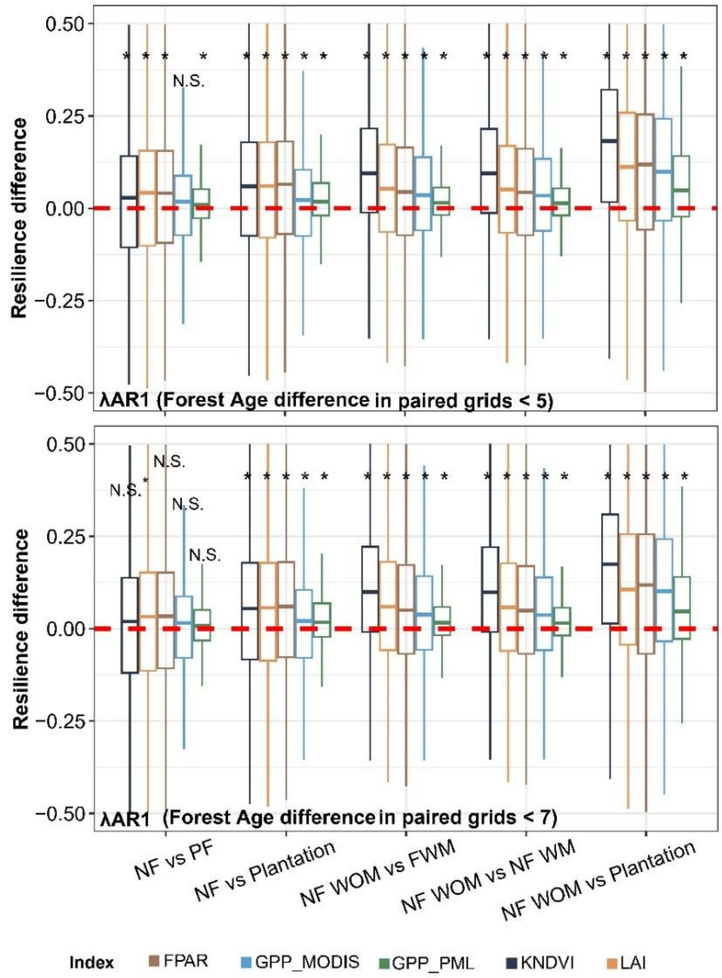


Figure S19. The impacts of forests' age difference on our main findings. Similar to Figure 2 (10 years), but with different thresholds of forests age difference (5 and 7 years).

2.Furthermore, in the final chapter discussing potential limitations, the authors should expand on the limitations of the CSD theory.

R: We agree with the importance of elaborating on the limitations of the critical slowing down (CSD) theory, as you suggested. **Action taken.** We have expanded on this point in the revised *Discussion Section*, highlighting the challenges of applying CSD in real-world ecosystems, such as (a) the influence of environmental noise, (b) the assumptions of stationarity, (c) the difficulty in detecting the time of resilience shifts, (d) and the impacts of the optical index on the accurate ecosystem resilience estimating. Here are the modified texts as follows: (Lines 492~513)

"It is important to acknowledge the potential limitations of CSD. Based on the CSD theory, changes in CSD metrics like λAR1 serve only as early warning indicators for resilience recovery and loss[15,

[67]. These indicators exert challenges to accurately pinpoint the time of critical transitions in system states. Furthermore, CSD metrics like variance are easily influenced by environmental noise, which can lead to inaccurate detection of slowing-down events[43]. For example, previous studies reported the critical slowing down in the Amazon Forest[20, 32], while the variance metric shows a decrease[44]. CSD theory assumes that systems exhibit stationary behavior. However, in ecosystems, external drivers such as abrupt events and human disturbances often lead to non-stationary conditions, which can weaken the accuracy of early warning signals derived from CSD[15]. Moreover, CSD indicators may be scale-dependent because ecological processes function across multiple spatial and temporal scales[45]. Ignoring this dependence may result in the uncertainties of resilience estimation. Additionally, optical remote sensing indices may affect the accuracy of resilience estimates. Since greenness indices such as the LAI are susceptible to saturation, assessing resilience using these indices in areas with high forest coverage may induce uncertainty[45, 68, 69]. It is thus recommended to use biomass indices such as GPP or VOD instead. Due to the coarse resolution of VOD indices (approximately 0.25°), we use the high-resolution GPP index to validate our findings, which show similar patterns to the results by optical remote sensing indices (more details in *Supplementary materials*). These limitations highlight the importance of careful application and interpretation of CSD-based resilience assessments."

3.Aside from these points, I have only a few minor comments detailed below:

Abstract: The abstract lacks specific mention of the aspect of resilience studied. I suggest mentioning the CSD theory in the abstract.

Introduction: The introduction is well-prepared, effectively outlining the concept of resilience, identifying the research gap, and clearly stating the study's objectives.

Methods: The authors utilized validated datasets, such as GFMD, the TerraClimate database, and MODIS. Although the methods involve some subjective choices, the sensitivity analysis provides reassurance about the robustness of the results. The classification of resilience into stable, increasing, or decreasing is based on a well-established theory and is therefore justified.

R: We appreciate your positive comments on the *Introduction and Method Section*. **Action taken.** In response to the comment on the *Abstract Section*, we have made the following revisions to address your concerns: (a) Mention of CSD Theory, (b) Clarification of Resilience Quantification, and (c) Clarification of method and datasets in the abstract. (Lines 4~22)

"Policymakers are concerned that newly planted trees may not provide ecosystem services as effectively as native forests. The efficient functioning of ecosystem services relies on the capacity of forest ecosystems to recover from external disturbances (i.e., ecosystem resilience). This study utilizes the critical slowing down (CSD) theory to quantify resilience differences among forest management types, particularly between planted forests (PF) and natural forests (NF). CSD resilience indicators such as AR1, λAR1, and λVar are estimated from satellite-based vegetation indices. This study further investigates the biophysical drivers for the resilience difference between PF and NF by machine learning. Results show that most forest management activities weaken forest resilience. Forest resilience experiences greater losses under intensive managements However, some forest managements can enhance resilience under specific conditions, such as wet climates. PF display higher resilience than NF in P/PET (the ratio of precipitation to potential evapotranspiration) > 1.5 (wetter climates), while NF is more resilient in P/PET < 1.5 (drier climates). We further detect such resilience differences in all climate zones between PF and NF decreasing from 2001 to 2015. These differences are influenced by biophysical contexts such as local climate, vegetation, and soil. Regions characterized by cold, wet conditions, dense vegetation, and higher soil fertility minimize differences in resilience, suggesting that rational site selection for afforestation can enhance forest resilience compared to native forests."

4.Detailed Comments:

Line 106: The phrase "In Spatial" is unclear—please clarify.

R: **Action taken.** To clarify, we intended to convey that the analysis results are based on spatial patterns. We have revised the text to improve clarity and readability. The modified text now reads (Lines 129~131): "Results show that areas characterized by human activity, as reflected in the forest management map, generally have lower resilience than natural forest without management (NF WOM) (Fig. 1)."

Line 108: Introduce the abbreviations GHFC and GFMD here (or in the Introduction) as they are used for the first time.

R: **Action taken.** We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The revised text now reads as (Lines 116~119):

"We derived the reference distribution of forest managements for the year 2000 (Fig. 1a) by overlapping tree cover data available for the year 2000 derived from the Global Hansen Forest Change Map (GHFC) and the spatial mapping of forest management types retrieved from the Global Forest Management Datasets (GFMD) for the year 2015."

In addition, we carefully reviewed the manuscript for similar instances where abbreviations were introduced without prior definition and made the necessary changes to ensure consistency and clarity (like AR1-LAI).

Line 176: Could you provide an explanation for this pattern already here? Why are the results different in these regions?

R: Thank you for your important question regarding the observed differences in resilience between PF and NF under varying P/PET (precipitation to potential evapotranspiration) conditions. Our findings suggest that these differences are closely tied to the sensitivity of vegetation to water and energy availability, which vary across dry and wet regions. In dry regions, where forest greenness is primarily limited by water availability, PF exhibit higher sensitivity to changes in water resources (Fig. 5). This means that PF are more vulnerable to disturbances such as droughts, which reduce soil moisture and water availability, leading to more severe degradation and slower recovery. In contrast, in wet regions, where forest greenness is mainly limited by energy availability, NF show greater responsiveness to variations in energy (Fig. 5). This difference in sensitivity to water and energy explains the observed resilience patterns between PF and NF across wet and dry regions (Figs. 3–5).

We hypothesize that the observed resilience differences may be related to the sensitivity of vegetation to water and energy availability. For example, if forests are more sensitive to water resources, disturbances such as droughts will result in more severe degradation due to the reduction in soil moisture and water availability. Consequently, the recovery process will be prolonged as the forest requires more time to accumulate resources and regenerate, especially when the regeneration rate is limited by drought-induced stress. Over time, this will lead to a decrease in the overall recovery rate of the forest ecosystem.

Additionally, we already discussed the reasons why are resilience different in these regions (Lines

427~448). Other factors may also contribute to this pattern: The species planted in PF are often selected for fast growth and may lack the physiological adaptations in some dry regions[27]. Additionally, in humid regions, management interventions in PF (e.g., thinning) might mitigate energy limitations more effectively than NF, enhancing PF resilience[28].

Action taken. We have added the sentence for clarity and to guide the reader to a detailed interpretation of the *Discussion Section*. The updated sentence now reads as (Lines 228~231): "This threshold likely reflects the point at which water availability becomes sufficient to alleviate stress from human managements, enabling PF to exhibit superior resilience to NF (more details in the *Discussion Section*)."

Line 183: Please check for grammatical accuracy—consider changing to "showing similar patterns" without "the."

R: **Action taken.** We have revised the sentence for improved accuracy and clarity. The updated sentence now reads as (Lines 139~142):

"Resilience estimated by using other vegetation indices (KNDVI, FAPAR, GPP_MODIS, GPP_PML) and alternative CSD indicators (λAR1, λVar) exhibits similar spatial patterns (Supplementary Fig. S2) and confirm the robustness of our assessment."

Line 242: Instead of "cycle," did you mean "circle"?

R: **Action taken.** We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The revised text now reads as (Lines 261~263):

"In the figure, a cross (×) indicates the mean value for each vegetation index, a circle (○) denotes the median value, and a red inverted triangle (▼) signifies the aggregated average of the means and medians."

Line 320: Please provide references to the "previous studies" that also documented a global decline in resilience.

R: **Action taken.** We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The revised text now reads as (Lines 348~349):

"This analysis reveals a general decline in the resilience of global forests, which is consistent with previous studies[18, 48]. ”

Line 425: I suggest adding a sensitivity analysis comparing younger NF and younger PN here. Additionally, consider mentioning the limitations of the CSD theory.

R: **Action taken.** We have addressed the related points in responses (above-mentioned), including a discussion of the limitations of the CSD theory (Lines 492~513).

Line 538: The reference "Simon et al." is missing from the reference list.

R: **Action taken.** We got the rules of referencing wrong and should have referenced the last name. We have corrected the text in the revised version of the manuscript. The revised text now reads as (Lines 690~691):

“The global forest age data are obtained from Besnard et al.[88], while elevation data were derived from the Shuttle Radar Topography Mission (SRTM) v4 dataset[89].”

Lines 614-615: Why didn't you use age as a predictor variable? Age could significantly impact resilience and potentially affect your findings. Please elaborate on this. Since PF are generally younger than NF, I recommend an additional sensitivity analysis comparing PF and NF of similar ages.

R: Thank you for your important suggestion regarding the potential impact of forest age on resilience and its role as a predictor variable. Below, we elaborate on our reasoning and the steps we have taken to address this concern:

Control for Forest Age in Pre-Processing:

In our analysis of resilience differences between planted forests (PF) and natural forests (NF), we controlled for large disparities in forest age by excluding paired grids where the age difference exceeded 10 years. This step minimizes the confounding effect of forest age on resilience differences. Similarly, we also controlled for elevation by excluding grids with an elevation difference exceeding 500 meters. These pre-processing steps ensure that the observed resilience differences are primarily attributable to forest management practices and other environmental factors, rather than large variations in forest age or elevation.

Additional Sensitivity Analysis:

In response to your suggestion, we conducted additional sensitivity analyses by applying varying age difference thresholds (e.g., 5, 7, and 10 years) to the paired grids (Supplementary Fig. S19). The results consistently showed that the observed patterns of resilience differences between PF and NF remained robust across these thresholds. This reinforces the reliability of our findings and indicates that forest age differences, within the tested thresholds, do not significantly alter the overall conclusions.

Rationale for Not Using Age as a Predictor Variable:

While forest age is undoubtedly an important factor, including it as a predictor variable in subsequent modeling would not be appropriate. Forest age was already accounted for in the pre-processing stage through paired grid filtering. Using it again as a predictor variable in the model would risk introducing redundancy and circular reasoning (i.e., independence), potentially biasing the results. Therefore, to maintain the independence of the model and avoid overfitting, we did not include forest age, canopy coverage, or elevation as predictor variables in our analysis.

Lines 617-621: Although you mention that "soil and vegetation variables" are detailed in the dataset availability section, I believe these should be briefly described in the Methods section, as they are critical inputs for your model.

R: **Action taken.** We agree that soil and vegetation variables are critical inputs for our model and should be briefly described in the *Method Section*. In response, we have revised the text to include the following details (Lines 735~740):

"The eight environmental factors include annual temperature (Temp derived from TeerraClimate), dryness index (P/PET derived from TeerraClimate), root depth (RD[92]), above-ground biomass (AGB[93]), below-ground biomass (BGB[93]), clay fraction (Clay[94]), sand fraction (Sand[94]), and cation exchange capacity (CEC[94]). These factors are selected for their known influence on forest resilience, including both climatic and land-surface variables."

Line 650: I couldn't find the reference to "Taylor et al."—please ensure it's included in the reference list.

R: **Action taken.** We have corrected the error by replacing the incorrect citation with the appropriate reference. The revised sentence now reads (Lines 767~769):

"Smith et al. calculated resilience by using all available records from a single satellite[45], while Forzieri et al. used growing season data from a single satellite[18]."

Additionally, we have carefully reviewed the entire manuscript to ensure that all in-text citations correspond accurately to the references listed in the bibliography.

I have not reviewed the code, but I have confirmed that it is available online as stated by the authors.

R: Thank you for confirming the availability of the code online. As part of our commitment to transparency and reproducibility, we will further organize and document the code after the manuscript is accepted. The finalized version of the code will be permanently archived on Fig share, ensuring long-term accessibility for the research community.

Response to Reviewer #4

Reviewer #4 (Remarks to the Author):

Review report

1. The study is focused on the forest management practices (human factors) and the (natural factors) driving forest ecosystem resilience. There are sustainable forest management practices that support forest ecosystem resilience while other unsustainable practices undermine forest ecosystem resilience. This is a very interesting topic, especially as the resilience of forest ecosystems is critical to the supply if ecosystem services benefiting people. The effect of natural factors that also influence forest ecosystem resilience, and that is a key objective of the manuscript is not capture in the manuscript title? The resilience of forest ecosystems is not only explained by forest management practices (type and intensity) but also more so, by biophysical factors, including climate change, and as such should be captured in the title.

Biophysical factors affected the impacts of forest management on ecosystem resilience?

R: Thank you for your important suggestion and constructive feedback. We agree that forest ecosystem resilience and the interplay between human activities and climate variables (i.e., biophysics factors). This study has indeed examined the influence of biophysical factors on the resilience differences between planted forests (PF) and natural forests (NF), which reflects the impact of forest management on resilience. For example, we found that regions with higher P/PET (precipitation to potential evapotranspiration ratio) show reduced resilience differences between PF and NF. It indicates that afforestation in wet zones may minimal the resilience gap between planted forests and local natural forests.

Action taken. Below, we address your concerns regarding the manuscript title. To address this, we propose revising the manuscript title to better capture the dual focus on forest management practices and biophysical factors. By combing your advances and the discussion with the author group, the modified new title now reads as: “Biophysical factors determining the effects of forest management on ecosystem resilience”. If you have further suggestions, let us know.

Additionally, we have made revisions throughout the manuscript to further emphasize the biophysical factors (i.e., we changed the original “context/background factors” to “biophysical factors” or “biophysical background/context”), such as Abstract (Lines 19~20), Introduction (Lines 98~99 and Lines 109~111), Figures’ caption (Line 254, Line 294), and Discussion (Lines 411~415).

Here is the related reversion:

In Abstract and Introduction Section:

"These differences are influenced by biophysical contexts such as local climate, vegetation, and soil." (Lines 19~20)

"(2) explore the associated biophysical factors (e.g., climate, soil, and vegetation conditions) affecting resilience differences between PF and NF by machine learning" (Lines 109~111)

In Result Section:

"Fig. 3 Relationships Between Resilience Differences (PF vs. NF) and Biophysical Factors." (Line 254)

"Fig. 4 Analyzing the Impact of Biophysical Factors on Vegetation Resilience Differences Between PF and NF." (Lines 294~295)

"The biophysical contexts including climate, vegetation, and soil variables may affect the resilience difference between PF and NF (details and models for other paired types in *Discussion*). We further use biophysical factors (and factors difference between PF and NF) to predict the resilience difference between PF and NF by the Extreme Gradient Boosting model (XGBT) (Fig. 4)." (Lines 265~269)

In Discussion Section:

"Additionally, aridity is not the sole biophysical factor influencing the resilience differences between PF and NF. Lower temperatures, higher aboveground and belowground biomass, and higher cation exchange capacity (e.g., indicating higher fertility) all contribute to reducing the resilience gap (Fig. 4)." (Lines 411~415)

2. Besides, I have some general major concerns.

Comparing planted forests and natural forests seems odd for me, as the difference in their resilience is quite obvious from their management practices. The results stated in the abstract are quite obvious: For example, "forest resilience suffers greater losses under intensive management, as seen in human plantations like agroforestry, which are much less resilient than unmanaged natural forests". And "Such resilience difference between PF and NF is related to background conditions such as local climate, vegetation, and soil". Topography also plays a role?

R: We respond to the reviewer's comment by addressing separately the two questions.

Question 1: why do we compare the resilience difference between PF and NF?

Results show that planted forests can surpass the resilience of natural forests. We indicated that natural forests include both managed and unmanaged ones. So, under certain conditions, the resilience of planted forests can exceed that of natural forests. Forest management can cause large damage to the resilience of natural forests without management (i.e., intact forests) (Fig .2). In addition, PF here are defined as trees with a longer rotation period (over 15 years) and for restoration purposes. Considering these reasons, the resilience of planted forests may indeed exceed that of natural forests. We also explain the mechanism from the perspective of climate sensitivity (see Discussion Section, Lines 427~448).

Based on previous studies and our findings, it seems difficult for planted forests to surpass the resilience of intact forests (Fig. 2). Our research indeed shows that unmanaged forests (i.e., natural forest without management) possess the highest forest resilience due to their high biodiversity and complexity (Fig. 2).

In the *Abstract Section*, such a description is just an example of "Forest resilience experiences greater losses under intensive management." Perhaps a more appropriate example here would be "as seen in human plantations like agroforestry, which are much less resilient than planted forests for restoration." After a discussion with our authors, we have finally deleted the sentence.

Additionally, we further describe the purpose of this study. Our research differs from that of Forzieri et al.[15], as we further detail the differences in resilience across different forest management types (not only focusing on the intact and non-intact forests). We found that forests subjected to intensive human managements such as rotation and agroforestry exhibit significantly reduced resilience. However, forests planted with the goal of restoration (i.e., PF with longer rotating years) do not necessarily have lower resilience than natural forests; this depends on their biophysical contexts such as climate, vegetation, and soil factors. Our results also indicated the regions where forest management has a lesser impact on resilience, serving as a policy reminder.

Action taken. We added the reasons why we mainly focus on the resilience between PF and NF (Lines 708~713).

“We first explored the biophysical drivers for resilience differences between PF and NF. This is particularly important given the widespread attention to restoration activities[21, 22, 23], such as the UN Decade on Ecosystem Restoration (2021–2030). Understanding the resilience differences between

PF and NF can provide valuable guidance for designing and implementing effective restoration strategies to enhance ecosystem resilience."

Question 2: Topography also plays a role?

Thanks for highlighting the potential role of topography. The topography is including factors such as landforms, elevation, slope, and aspect. In this study, we have taken steps to mitigate the impact of elevation on the estimation of resilience differences. Specifically, to control for the potential confounding effects of varying forest age and elevation, we ensured that the differences in these variables between paired forests did not exceed 10 years and 500 meters, respectively. Here are the words from the original texts in Lines 687~691 (*Method Section*): "To mitigate the potential effects of varying forest age and elevation on vegetation resilience, we ensured that the differences in these variables within each paired grid were restricted to within 10 years and 500 meters, respectively[59]. The global forest age data are obtained from Besnard et al.[88], while elevation data were derived from the Shuttle Radar Topography Mission (SRTM) v4 dataset[89]."

Action taken. In the *Result Section*, we re-emphasize this point. (Lines 159~165)

"Specifically, grids are paired using the sliding window method, which moves a fixed-size window across the study area. Within each window, paired grids are selected based on similar forest age (differences of ≤10 years), elevation (differences of ≤500 meters), and a minimum of 20% forest coverage for each management type (more details in the *Method Section*). We point out that the effects of potential covariates have been ruled out within the sliding window approach."

Additionally, while we accounted for elevation differences in our analysis, we acknowledge that broader topographical features—such as slope, aspect, and landforms—may also influence resilience differences between PF and NF on a global scale. These factors can affect soil water availability, microclimate, and vegetation growth, potentially interacting with forest management practices. In response to your comment, we modified and added related limitations in the discussion, as follows: (Lines 514~517)

"This study prominently examines the differences in resilience between PF and NF, further focusing on their response to biophysical factors. Certain factors (e.g., landforms and slope), that influence the vegetation recovery process, may not have been adequately accounted for in the model[70]."

3. These are obvious well documented results and do not provide any novelty, even where a quantified difference is provided. Is the resilience of well managed planted forests at the same level as or even more than unmanaged natural forests?

My suggestion is to compare similar forest ecosystem types (e.g., managed and unmanaged natural forests; planted commercial forests and planted forests mainly aimed to restore ecosystem functions and service, etc.). Comparing similar forest ecosystem types with different or similar forest management practices and biophysical conditions would be more intriguing in understanding differences in their resilience and inform restoration and conservation decision making. Comparing within planted forests with primary objective for restoration functions and services and planted forests under the commercial forestry approach can enhance our understanding of the resilience of these ecosystems in terms of supply of ecosystem services. The resilience of planted forests monocultures and multiple tree species forests can be compared to further understand why forest monocultures are discouraged.

R: We agree with your comments that comparing the resilience of similar forest types is essential. We added the additional analysis to compare the resilience difference between similar forest types, including managed versus unmanaged natural forests and long-rotation versus short-rotation planted forests (i.e., planted commercial forests and planted forests mainly aimed to restore ecosystem functions) (Fig S24-S29). In addition, we investigate the relationship between tree species richness and resilience, to further validate why forest monocultures are discouraged (Fig S24). Here, we further respond to the reviewer's comment by addressing separately the two questions.

Question 1: "Is the resilience of well-managed planted forests at the same level as or even more than unmanaged natural forests?"

Although we responded to this question briefly in the previous comment, here we will express it in detail. After many pairwise comparisons, our findings indicate that unmanaged natural forests exhibit the highest resilience among all forest management types potentially due to their complex and integrated ecological process (Fig. 2 and Supplementary Fig. S26). Previous studies have suggested that management practices such as irrigation, rotation, and logging reduce ecosystem complexity and biodiversity, thereby diminishing the forests' capacity to recover from disturbances. Our findings also indicated that greater intensity of forest management corresponds to reduced resilience. For example, the resilience difference between plantations and NF is larger than that between planted forests and NF (Fig. 2). The plantations include the planted forests, planted forests with rotating up to 15 years, agroforestry, and oil palms, which have more intense human activities or management. Through this comparison, we can further infer those commercial forests (such as timber forests-PF with short rotating years, agroforestry, and oil palms), which are more intensively managed, have lower resilience than planted forests aimed at restoration. Thus, increased management does not make planted forests more resilient than unmanaged natural forests (even makes planted forests less resilient).

However, under specific biophysical conditions (e.g., wet conditions with P/PET > 1.5), planted forests (PF) can exhibit higher resilience than natural forests, which both include the managed and unmanaged NF. These differences hinge on various biophysical factors including climate, soil, and vegetation, which vary significantly across regions. This comparison and investigations inform optimal reforestation activities for reducing the resilience between PF and NF. Furthermore, although planted forests can surpass natural forests in resilience under certain conditions (Fig. 2-5), the resilience of unmanaged natural forests is still larger than planted forests (Supplementary Fig. S26). This further implies that human managements may significantly impair the resilience of natural forests.

In summary, it is important to distinguish between three definitions: well-managed forests, intensively managed forests, and human-managed forests considering the biophysical contexts. "Well-managed" represents the result of human intervention that is challenging to quantify precisely (it's multi-dimensional including environment, economy, and benefits). Our findings only suggest that human managements can generally lead to a reduction in forest resilience. Forests under higher levels of human pressure (i.e., intensive managements) tend to exhibit lower resilience. However, when biophysical contexts are considered in managements (e.g., choosing humid or cold regions for afforestation), planted forests maybe have higher resilience than natural forests. However, a noticeable resilience gap remains compared to unmanaged natural forests. Natural forests without management have much higher resilience than that in natural forests with management (Fig. 2).

Question 2: why not compare similar forest ecosystem types? (compare resilience between planted commercial forests and planted forests mainly aimed to restore ecosystem functions and services)

We agree with your suggestions, and we first added the analysis and discussion of comparing resilience between planted commercial forests and planted forests mainly aimed at restoring ecosystem functions and services. In the original analysis (Fig. 2), we conducted two comparisons: the resilience difference between plantations and natural forests, and the resilience difference between planted forests and natural forests. The plantations conclude the planted forests, planted forests with rotating up to 15 years, agroforestry, and oil palms. We found that the resilience difference between plantations and natural forests (NF) is larger than that between planted forests and NF. This further partly supports that commercially oriented plantations have lower resilience compared to restoration-oriented planted forests.

Furthermore, we attempted to pair the grids method to directly compare the resilience between wood-oriented plantations (refer to the planted forests up to 15 rotating years) and restoration-function-oriented plantations (refer to the planted forests with over 15 rotating years). Unfortunately, the number of paired grids was limited (only ten to several dozen), which may compromise the robustness of our analysis. We have to consider the number of paired grids to design paired experiments. In this regard, we reclassified the GFMD to ensure that there were enough paired grids for use in the analysis. In other words, the comparison results of different forest management types in Figure 2 already consider the number of paired grids available. Other pairing results may not have sufficient grids for further analysis (typically requiring more than 500 grids).

Action taken.

(1) In response, we added the limitation of the paired grids method in the *Method Section* (Lines 700~706):

"Therefore, we designed six types of comparative experiments to examine the differences in resilience across different vegetation management types, including resilience between NF and PF, NF and Plantations, NF WOM and F WM, NF WOM and NF WM, and NF WOM and Plantations. For each type of comparison, over 500 paired grids were analyzed to ensure statistical robustness and comprehensive coverage of the different management comparision."

(2) To provide a more robust analysis, we utilized original crowdsourced data to examine the resilience of wood-oriented plantations and restoration-function-oriented plantations (Supplementary Fig. S25). By definition[29], planted forests with shorter rotation periods (<15 years) are often commercially oriented plantations aimed at timber production, whereas those with longer rotation periods are typically restoration-oriented planted forests (> 15 years). Our enhanced

comparison revealed that: In regions with P/PET < 3.5, planted forests with low rotation years (i.e., wood plantation) exhibit significantly lower resilience compared to those with high rotation years (i.e., planted forest for restoration). In regions with P/PET > 3.5, the difference in resilience between long and short-rotation age plantations becomes minimal. Overall, the resilience of long-rotating planted forests is larger than that with short rotation periods. We further found that the resilience of oil palm plantations is generally lower than that of long-rotation planted forests, again highlighting that commercially oriented plantations have lower resilience compared to restoration-oriented planted forests.

In addition, we further investigated the relation among P/PET, tree species richness, and forest resilience. We found that resilience in all plantations increases with higher tree species richness. The increase in tree species richness further enhances resilience under similar climate conditions (Fig. S26). These findings highlight the importance of encouraging multi-species plantations as a policy recommendation for the enhancement of forest resilience and ecosystem services.

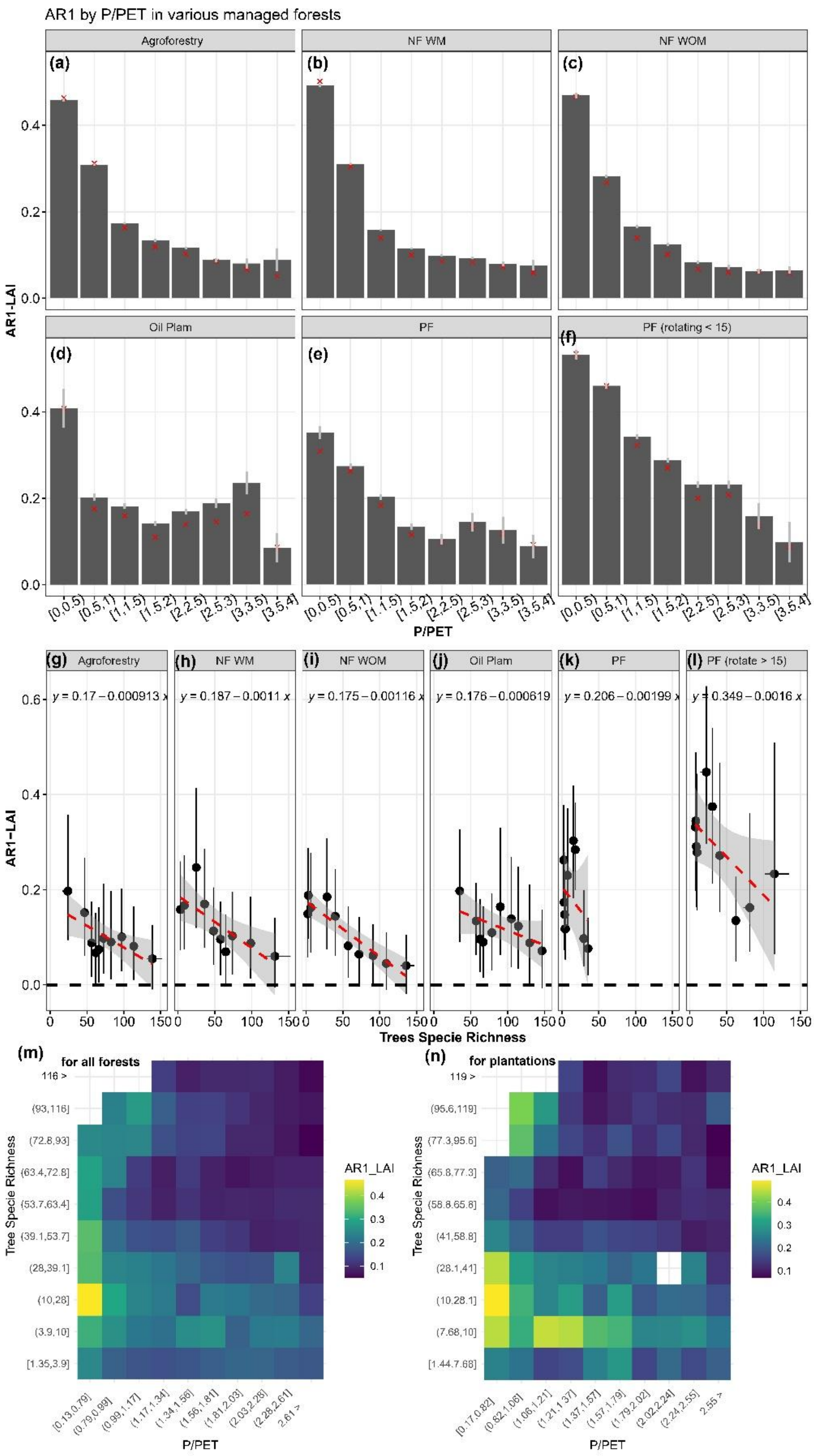


Figure S26.The relation among forest resilience, P/PET, and tree species richness across various managed types of forests. (a-f) The AR1 by P/PET in various forest management trees. (g-l) The relation between AR1-LAI and tree species richness across various management forests. (m-n) The AR1-LAI heatmaps illustrate the relationship between tree species richness and P/PET across different forest types. Panel (m) shows the AR1_LAI for all forests, while panel (n) focuses on plantations. Here we used the original crowdsourcing datasets, long-term LAI resilience, tree species richness, and the P/PET for statistics. The grey bar is the mean value, and the error bar is the mean ± 2*standard error. The red cross is the median value. All datasets such as forest resilience, P/PET, and tree species richness are resampled to 1km. Global tree species richness has 0.025° spatial resolution, derived from https://figshare.com/articles/dataset/The_global_map_of_tree_species_richness/17232491.

According to the reviewer's suggestion, we added above-mentioned discussion of resilience comparison between planted commercial forests and planted forests mainly aimed at restoring the ecosystem. We also added the description of afforestation by multiple tree species forests. Here is the related discussion, now reading as (Lines 472~482):

"For example, PF with rotations up to 15 years (i.e., wood plantations/shorter-rotation PF) exhibit lower resilience compared to those with rotation periods over 15 years (i.e., planted forests for restoration/longer-rotation PF) (Supplementary Fig. S26). These shorter-rotation PF are economically composed of a single species and are often partially harvested at regular intervals[66]. As a result, species richness and biodiversity in shorter-rotation PF are generally lower compared to longer-rotation PF, reducing the capacity to recover from disturbances[52]. Our study also supports the indications that the increase in tree species richness can further enhance forest resilience (Supplementary Fig. S26). It encourages multi-species plantations as a policy recommendation for the enhancement of forest resilience for afforestation."

Comparing managed and unmanaged natural forests can provided important insight on the effect of forest silvicultural management practices on forest ecosystem resilience. Unmanaged old growth forest may sequester less carbon due to older age trees compared to young secondary regrowth forests. Do unmanaged natural forests need some sought of management intervention to increase resilience or should they be kept intact?

The resilience of forest ecosystems is not only explained by forest management practices (type and intensity) but also more so, by biophysical factors. But it is clear this was not investigated in this study, making the scope so one dimensional. This is clearly acknowledged by the authors. For example:

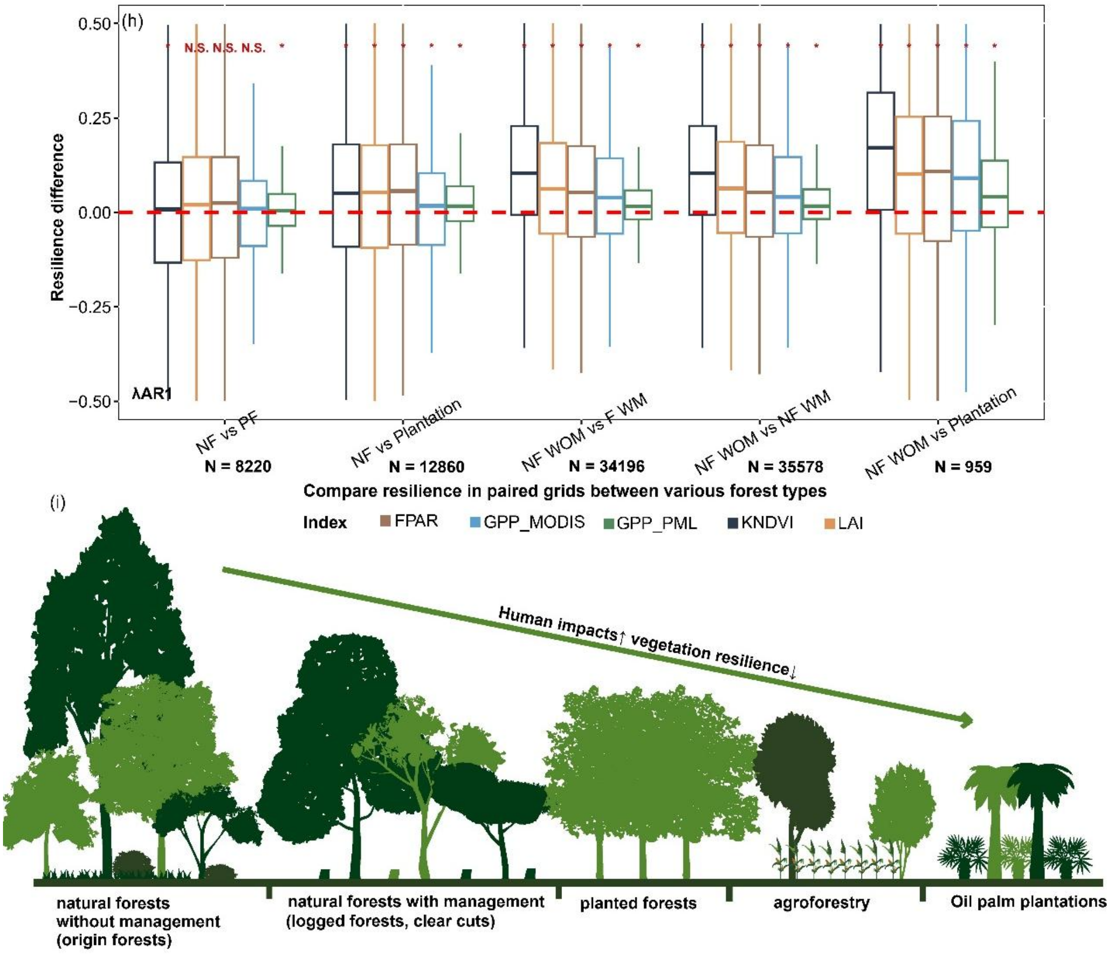
In line 446 to 448, the authors clearly stated that "…the narrowing disparity in resilience between PF and NF may be attributed to multiple factors, including vegetation physiology, climate change, and the intensity of human activities. This complex issue merits further investigation".

In lines 459 to 460, "… but there is less exploration of whether the climate, soil, and vegetation background can mitigate the effects of other forest managements".

R: We agree with comments on the importance of comparing forest ecosystems under similar management practices. According to the reviewer's comment, we built additional four XGBT models for investigation of the mitigation of the effects of forest management by the climate, soil, and vegetation background (Supplementary Fig. S29). Here, we respond to the reviewer's comment by addressing separately the two questions.

Question 1: Do unmanaged natural forests need some sought of management intervention to increase resilience or should they be kept intact?

We have employed a paired study design to compare: (1) Managed Natural Forests vs. Unmanaged Natural Forests and (2) Managed Forests vs. Unmanaged Forests. The results indicate that managed natural forests exhibit significantly lower resilience compared to unmanaged natural forests, highlighting the potential adverse effects of management practices on natural forest ecosystems (Part of Figure 2). The results of our crowdsourced analysis also show similar patterns (Supplementary Fig. S26-S27). Thus, for unmanaged natural forests, maintaining their integrity (i.e., being kept intact) can enhance their resilience, aligning with previous study[30]. Unmanaged natural forests have the highest resilience among all forest management types. We should avoid the disruption of pristine ecological processes by human management such as irrigation changing the root zone or trimming reducing biomass[31]. This disruption could reduce the resilience of unmanaged natural forests. Policymakers are designating protected areas to facilitate natural regeneration processes.



Part of Figure2. Further comparisons of resilience are performed across various forest management types using the same methodologies. These comparisons include NF vs. PF, NF vs. Plantations, NF WOM vs. F WM, NF WOM vs. NF WM, and NF WOM vs. Plantations. The notation "X vs. Y" indicates the resilience difference, calculated as the resilience of Y minus that of X. Statistical significance tests are performed to assess whether these resilience differences are statistically distinguishable from zero. Bars marked with an asterisk (*) denote statistical significance at $p < 0.05$, while bars labeled with 'NS' denote non-significant differences. (l) A conceptual figure illustrates different forest management types (e.g., NF WOM). Abbreviations used in the figure include: NF-Natural Forests, PF-Planted Forests, NF WOM-Natural Forests without Managements, NF WM-Natural Forests with Managements, and F WM-Forests with Managements.

Question 2: What's the relationship between resilience differences and biophysical factors?

We agree that these biophysical factors play a significant role in resilience differences, and we did not fully investigation in original main texts. We have further developed and integrated relevant models to offer a more comprehensive explanation of our results.

Action taken. We further built an XGBT model to explore other comparing resilience between NF and Plantations, NF WOM (without management) and F WM (with management), NF WOM and NF WM, and NF WOM and Plantations. These models allow us to explore the role of biophysical factors, such as tree root depth, climate, and soil conditions, in shaping forest resilience under different management practices. For instance, we found that the depth root of trees is the most important factor for predicting the resilience differences between managed and unmanaged NF. Therefore, we should avoid management activities like irrigation in areas with deep roots, as this significantly disrupts the development of natural forest root zones, reducing their ability to mitigate disturbances by accessing groundwater (thus decreasing resilience). Previous study has also reported that deep groundwater roots can alleviate the extent of damage to the Amazon rainforest during droughts[32], leading the vegetation recovery soon. Here are the modified texts and figures:

In Method Section: (Lines 745~748)

"In addition, we also established XGBT models to explore other comparisons of resilience between NF and Plantations, NF WOM and F WM, NF WOM and NF WM, and NF WOM and Plantations (more details in *Supplementary materials*) (Supplementary Fig. S29)."

In Discussion Section: (Lines 514~533)

"This study prominently examines the differences in resilience between PF and NF, further focusing on their response to biophysical factors. Certain factors (e.g., landforms and slope), that influence the vegetation recovery process, may not have been adequately accounted for in the model[70]. Furthermore, we also built XGBT models to investigate the influence of biophysical contexts on the resilience differences between NF and Plantations, NF WM and NF WOM, NF WOM and F WM, and between NF WOM and Plantations (Supplementary Fig. S29). Results reveal that temperature is the most important factor in predicting the resilience differences between NF and Plantations, and between NF WOM and Plantations. The root depth is the most important factor in driving resilience differences between NF WM and WF WOM (or between F WM and NF WOM). Thus, we should avoid management activities like irrigation in regions with deep-rooted forests, as such activities disrupt the development of natural forests' root zones, reducing their ability to mitigate disturbances by accessing groundwater. A previous study reported that deep-rooted forests can alleviate drought-induced damage in the Amazon rainforest[71], leading to faster recovery. More detailed analyses and discussions regarding the XGBT model results can be found in the *Supplementary Materials*. Ultimately, our comparisons across different forest management types offer crucial guidance for sustainable forest development. Policymakers can better choose where to perform managements to mitigate its potentially negative effects on resilience."

In Supplementary Materials: (Lines 91~114)

"**Supplementary Note 4: The relation between biophysics factors and the resilience difference among various managed forests.**

We further developed four XGBT models to explore the relationships between resilience differences and environmental factors (Supplementary Fig. S29). Specifically, the models analyze resilience differences between NF and Plantations, NF WOM and F WM, NF WOM and NF WM, and NF WOM and Plantations. The overall performance of these models, as measured by the *KGE* (Kling-Gupta Efficiency), ranged between 0.6 and 0.7 (Supplementary Fig. S29), indicating good performance. Results reveal that higher temperature, the most important variable, may exacerbate the resilience differences between NF and Plantations, and between NF WOM and Plantations (Supplementary Fig. S29a, d). In regions with high temperatures (e.g., tropical regions), Plantations are often dominated by monoculture cash crops, such as oil palm, which lack the structural and biological complexity compared to local forests[9]. Furthermore, human activities like logging prominently disrupt the integrity and biodiversity of forests, leading to lower resilience in Plantations compared to NF WOM or NF WM. Furthermore, management practices in regions with deeper root depth may increase the resilience differences between NF WM and NF WOM (or F WM and NF WOM) (Supplementary Fig. S29b, c). Forests with deep-rooted depth often rely on groundwater, which provides a buffer against drought or water stress[14]. However, human managements can disrupt this reliance through irrigation[15], making forests more sensitive to drought or water variability, thereby significantly reducing resilience. Furthermore, in regions with deep root depth (e.g. mountainous regions), root and vegetation structures in NF WOM are often more complex and diverse. Human management in such regions tends to simplify root systems, degrade soil quality, and reduce biodiversity/complexity[16]. Overall, to mitigate the negative impacts of human forest management on resilience, it is essential to implement managements considering climate conditions and vegetation biophysical characteristics (Fig. 4 and Supplementary Fig. S29)."

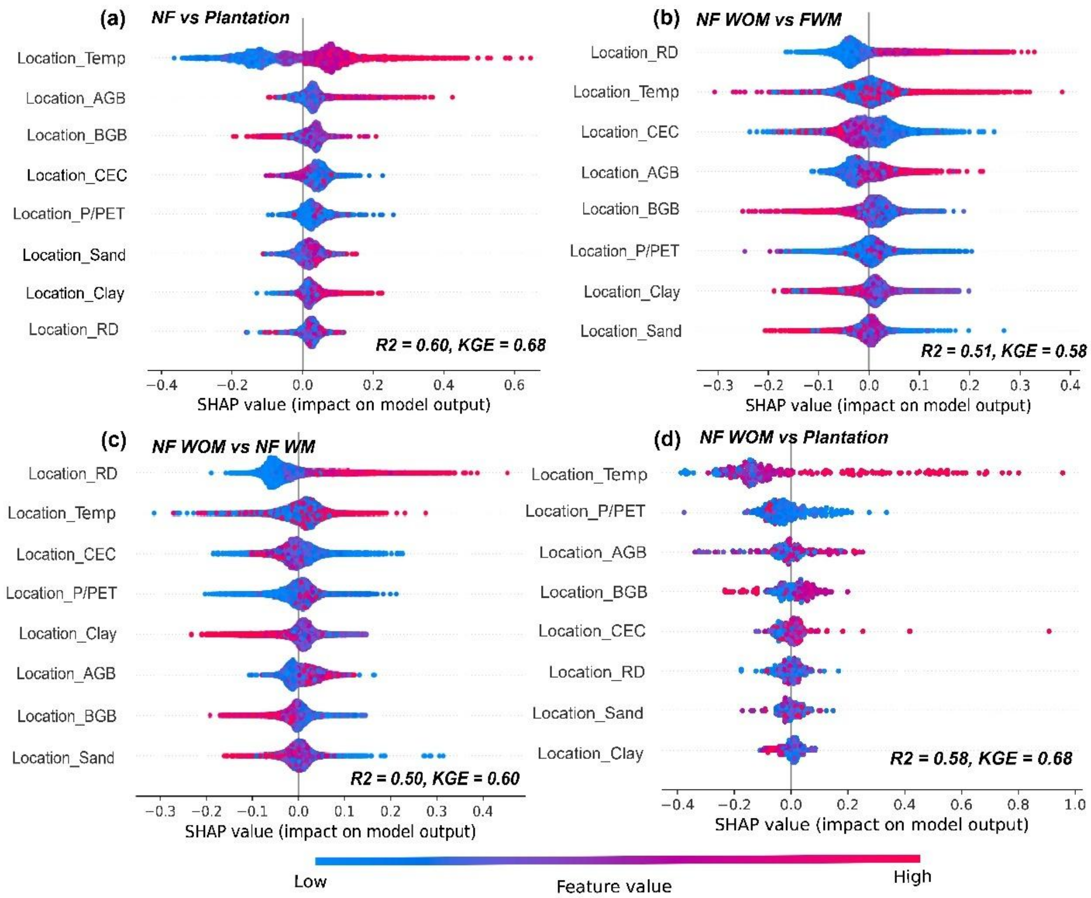


Figure S29. Similar to Figure 4, but focusing on predicting the resilience differences for the following comparisons: NF vs. Plantation, NF WOM vs. F WM, NF WOM vs. NF WM, and NF WOM vs. Plantation. The R2 and KGE in the lower right corner are derived from the 5-fold cross-validation.

References:

1. Dakos, V., Van Nes, E. H., d'Odorico, P. & Scheffer, M. Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology* **93**, 264–271 (2012).
2. Smith, T., Zotta, R.-M., Boulton, C. A., Lenton, T. M., Dorigo, W. & Boers, N. Reliability of resilience estimation based on multi-instrument time series. *Earth Syst. Dynam.* **14**, 173–183 (2023).
3. Smith, T., Traxl, D. & Boers, N. Empirical evidence for recent global shifts in vegetation resilience. *Nat. Clim. Change* **12**, 477–484 (2022).
4. Smith, T. & Boers, N. Reliability of vegetation resilience estimates depends on biomass density. *Nat. Ecol. Evol.* **7**, 1799–1808 (2023).
5. Zhang, Y. *et al.* Warming and disturbances affect Arctic-boreal vegetation resilience across northwestern North America. *Nat. Ecol. Evol.* **8**, 2265–2276 (2024).
6. Blaschke, L. L. *et al.* Spatial correlation increase in single-sensor satellite data reveals loss of Amazon rainforest resilience. *Earth's Future* **12**, e2023EF004040 (2024).
7. Boers, N. Observation-based early-warning signals for a collapse of the Atlantic Meridional Overturning Circulation. *Nat. Clim. Change* **11**, 680–688 (2021).
8. Pickering, M. et al. Assessing the relationship between forest structural diversity and resilience in a warming climate. EGU General Assembly 2024, Vienna, Austria, 14–19 Apr 2024, EGU24-16017 (2024).
9. van de Leemput, I. A. *et al.* Critical slowing down as early warning for the onset and termination of depression. *Proc. Natl. Acad. Sci. USA* **111**, 87–92 (2014).
10. Hu, Z., Dakos, V. & Rietkerk, M. Using functional indicators to detect state changes in terrestrial ecosystems. *Trends Ecol. Evol.* **37**, 1036–1045 (2022).
11. Dakos, V., Carpenter, S. R., van Nes, E. H. & Scheffer, M. Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philos. Trans. R. Soc. B* **370**, 20130263 (2015).
12. Hansen, M. C. *et al.* High-resolution global maps of 21st-century forest cover change. *Science* **342**, 850–853 (2013).
13. Camps-Valls, G. *et al.* A unified vegetation index for quantifying the terrestrial biosphere. *Sci. Adv.* **7**, eabc7447 (2021).
14. Wang, Q., Moreno-Martínez, Á., Izquierdo-Verdiguier, E., Camps-Valls, G. & Zhu, X. Estimation of vegetation traits with kernel NDVI. *ISPRS J. Photogramm. Remote Sens.* **195**, 408–417 (2023).
15. Forzieri, G., Dakos, V., McDowell, N. G., Ramdane, A. & Cescatti, A. Emerging signals of declining forest resilience under climate change. *Nature* **608**, 534–539 (2022).
16. Yu, Z., Liu, S., Wang, J., Wei, X., Schuler, J., Sun, P., Harper, R. & Zegre, N. Natural forests exhibit higher carbon sequestration and lower water consumption than planted forests in China. *Glob. Change Biol.* **25**, 68–77 (2019).
17. Besnard, S. et al. Mapping global forest age from forest inventories, biomass and climate data. *Earth Syst. Sci. Data* **13**, 4881–4896 (2021).
18. Jarvis, A., Reuter, H. I., Nelson, A. & Guevara, E. Hole-filled SRTM for the globe Version 4, available from the CGIAR-CSI SRTM 90m Database. (2008).
19. Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).

20. Liu, Y., Kumar, M., Katul, G. G. & Porporato, A. Reduced resilience as an early warning signal of forest mortality. *Nat. Clim. Change* **9**, 880–885 (2019).
21. Zhong, Z. et al. Higher sensitivity of planted forests' productivity than natural forests to droughts in China. *J. Geophys. Res. Biogeosci.* **126**, e2021JG006306 (2021).
22. Fernández, M.E., Gyenge, J. & Schlichter, T. J.T. Water flux and canopy conductance of natural versus planted forests in Patagonia, South America. *Tree Physiol.* **23**, 415–427 (2009).
23. Martín-Benito, D. et al. Response of climate-growth relationships and water use efficiency to thinning in a Pinus nigra afforestation. *For. Ecol. Manage.* **259**, 967–975 (2010).
24. Kuuluvainen, T. & Gauthier, S. Young and old forest in the boreal: critical stages of ecosystem dynamics and management under global change. *For. Ecosyst.* **5**, 1–15 (2018).
25. Bohn, F. J., May, F. & Huth, A. Species composition and forest structure explain the temperature sensitivity patterns of productivity in temperate forests. *Biogeosciences* **15**, 1795–1813 (2018).
26. Morin, X., Fahse, L., De Mazancourt, C., Scherer-Lorenzen, M. & Bugmann, H. Temporal stability in forest productivity increases with tree diversity due to asynchrony in species dynamics. *Ecol. Lett.* **17**, 1526–1535 (2014).
27. Brockerhoff, E.G. et al. Plantation forests and biodiversity: oxymoron or opportunity? *Biodiversity and Conservation.* **17**, 925–951 (2008).
28. Bauhus, J. et al. Mixed-species forests: ecology and management. *Mixed-species forests: ecology and management*. Springer; 2017. p. 1–24. https://doi.org/10.1007/978-3-662-54553-9_7
29. Lesiv, M. et al. Global forest management data for 2015 at a 100 m resolution. *Sci. Data* **9**, 199 (2022).
30. Liu, X. et al. Enhancing ecosystem productivity and stability with increasing canopy structural complexity in global forests. *Sci. Adv.* **10**, eadl1947 (2024).
31. Payn, T. et al. Changes in planted forests and future global implications. *Forest Ecol. Manag.* **352**, 57–67 (2015).
32. Chen, S. et al. Amazon forest biogeography predicts resilience and vulnerability to drought. *Nature* **631**, 111–117 (2024).

Response to Reviewers

The authors sincerely thank the anonymous reviewers for their valuable and constructive feedback. To present the content more clearly and in line with Reviewers' suggestions, we have polished and revised the manuscript. We have carefully addressed all concerns and revised the manuscript accordingly. Below, we provide detailed responses to each comment from the reviewers. In the revised manuscript, reviewer comments are shown in black, our revisions are highlighted in red, and our responses are indicated in blue. Please note that all line numbers refer to the revised version of the manuscript. Additionally, the references cited within the red-colored text are available in the tracked manuscript, while those in the blue responses are included in this document.

Response to Reviewer #1

Reviewer #1 (Remarks to the Author):

Overall impression

The authors have made a remarkable effort to make the manuscript more legible. They have effectively answered several concerns of the reviewers. However, now that it is clearer (notably on the metrics used), my overall feeling is that there is inherent flaws in this study. Unfortunately, these flaws confirm to me the first impression that it should not be accepted. I will elaborate below my main concern, and the expose some of the more specific issues I encountered while reading the document.

R: We sincerely thank the reviewer for your encouragement of our revisions in Round 1. We are also very grateful for your comments and suggestions, which have significantly improved our work. Additionally, we have been greatly inspired and motivated by your insightful feedback. The reviewer's time and effort on commenting on our work are highly appreciated as well.

We have carefully considered and thoroughly addressed the new concerns raised in revision 2, including the application of CSD theory to forest resilience, the definition of tipping points, the treatment of normalization, and the implementation of the MVC method.

The reviewer's main concern appears to focus on the application of CSD indicators (which is also related to the definition of resilience), particularly emphasizing the quantification of resilience requires clarifying the specific disturbance. The reviewer thinks that resilience is the recovery after major and specific disturbances (e.g., droughts or fires). It is the concept of engineering resilience. However, this is not the only definition of resilience. As supported by the literature[1–9], we have adopted the concept of ecological resilience, which is defined as the capacity of a system to remain in its current state despite external perturbations. This is mathematically equivalent to measuring the short-term response to continuous disturbances (including strong and weak disturbances), rather than specific events. Therefore, based on the existing literature and mathematical relationships, we selected these CSD-based indicators (such as AR1) to quantify ecological resilience. In addition, we also calculated the empirical response rate (based on the recovery curve after disturbances) and CSD resilience indicators based on crowdsourced data to verify the robustness of our results (Fig. R1c-f). The similar patterns between empirical response rate and CSD indicators ensure the robustness of our findings.

The reviewer also mainly expressed concerns regarding the use of remote sensing data within the CSD framework. In fact, recent studies[3,10–19] have estimated early warning indicators of resilience using bi-weekly or monthly vegetation datasets, and our study is based on this established approach. The resilience results of this study show a spatial pattern consistent with existing findings (Fig. R6), confirming the correctness and feasibility of our method. Furthermore, we emphasize that our study has validated the results across different temporal (8-days to monthly, Fig. S22~25) and spatial resolutions (500m to 1000m, Fig. S29), demonstrating the robustness of our findings. Detailed point-by-point responses to each specific issue can be found in the corresponding sections below.

Specifically:

General issues

The main problem I have with this study relates to the use of the CSD theory specifically on these global vegetation time series of remote-sensing derived indicators in view of characterizing alleged differences in resilience between natural and planted forests (i.e. the main goal of the study). As the authors state (L85) "According to CSD theory, as systems approach tipping points where system states change, their sensitivity to disturbances increases, and recovery from disturbances slows down, indicating a loss of resilience". To begin with, it would be necessary to state specifically how these concepts are specifically mapped to the study case at hand of forests.

R: Thanks for your important comments. Here we interpret the relationship between forests resilience and CSD (Critical Slowing Down) theory. **There are two major definitions of resilience**[10,11,20,21]. Forest resilience is usually defined as the capacity to recover after a disturbance. The disturbance is specific natural (e.g., wildfire, insect outbreaks, drought) or human-induced (e.g., logging, irrigation) event[22]. This definition corresponds to **engineering resilience**. The engineering resilience focuses on the system's response to the specific and strong external shocks[23]. For instance, some studies investigated the recovery of the greenness in California after suffering from the millennium drought from 2012 to 2016[24]. That is the classical definition of resilience, but it is not the only one. Resilience can also be defined as the ability of a system to remain in its original state (that is ecological resilience)[10]. The **ecological resilience** concerns the system's continuous response to a sequence of disturbances, both strong and weak (i.e., not limited to a specific disturbance)[12,25]. This study adopts the latter definition (ecological resilience), in line with the cited literature[10–12], as managed forests are subject to continuous human and natural disturbances.

In our analysis, forests (including PF and NF) are exposed to an ongoing mix of disturbances, including climatic variability and human management (i.e., strong and small disturbances). The way the system reacts (e.g., resist disturbances and maintain their state) to that sequence of disturbances constitutes its ecological resilience. We quantify this resilience by CSD indicators such as AR1, λAR1 and λVar. These indicators, in addition to quantifying the memory effect or fluctuation of the time-series data, are also closely related to the recovery rate after disturbance. Previous study[11] calculated the empirical recovery rate, which was estimated by the recovery curve of vegetation after each strong disturbance (Fig. R1a). **Furthermore, they discovered a close non-linear relation between AR1 and empirical recovery rate** (Fig. R1b). Thus, AR1 serves as a reliable indicator for resilience.

Furthermore, about the physical meaning of these indicators in CSD theory: CSD theory predicts that, as an ecological system loses resilience and approaches a tipping point (TP), its recovery from perturbations slows down—manifesting as "critical slowing down" [2]. This is detectable in time series as an increase in AR1 and variance, since disturbances leave longer-lasting traces ("memory") in the system[11]. The longer memory traces can be interpreted as the persistent effects of past disturbances. If the effects of past disturbances are more pronounced, it indicates that the system has not fully recovered from previous shocks, and its recovery capacity is low (AR1 is high)[26]. When the ecosystem is "healthy" and has high resilience, the effect of disturbances will dissipate quickly, and the correlation between adjacent observations in the time series is low (AR1 is low). Thus, AR1 indicator is now well established in the CSD literature as a reliable proxy for ecological resilience[2,3,6].

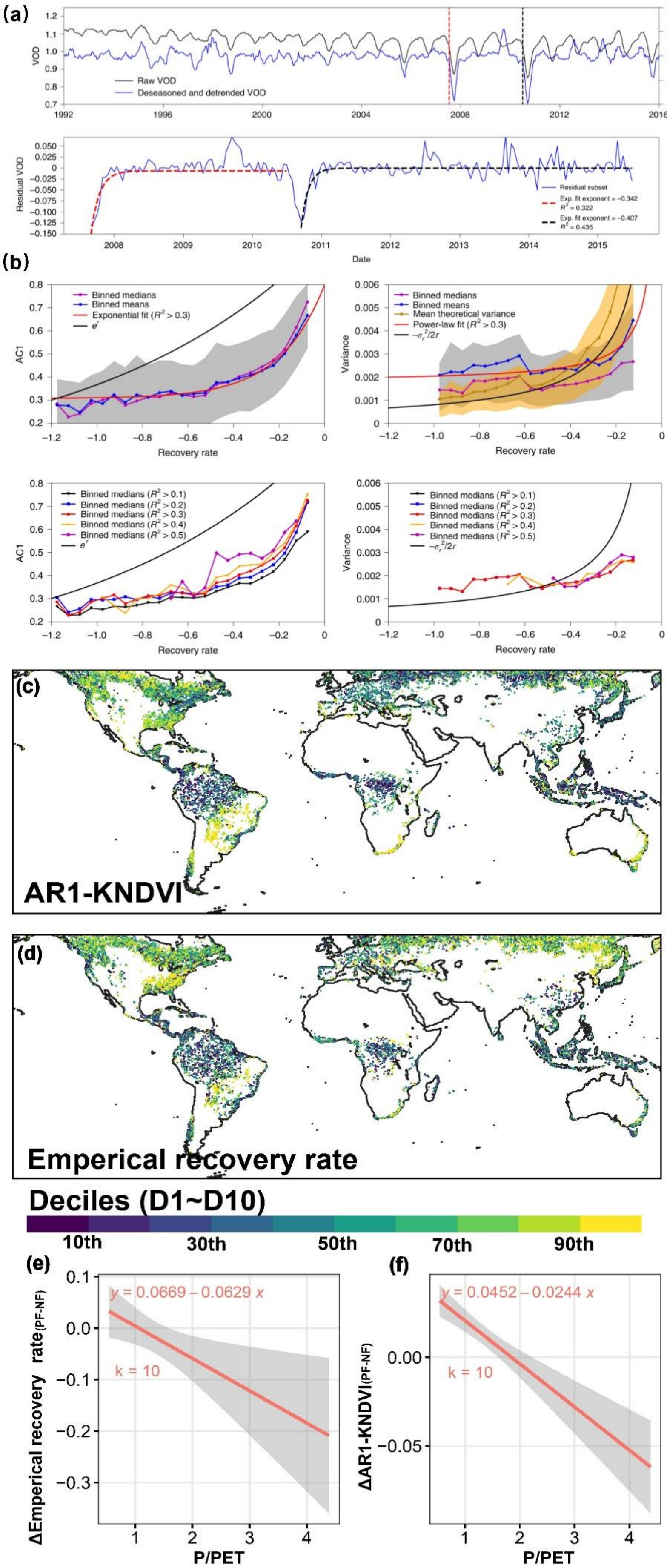
Moreover, there is a rigorous mathematical basis linking AR1 to the recovery process[26–28]**.** Specifically, the AR1 coefficient in a time series is inversely related to the system's recovery rate following a perturbation. This relationship is derived by mapping the discrete-time process to the continuous-time exponential decay model (equations in the Lines 328~355 of the supplementary materials), leading to the approximate equation $recovery = -\ln(AR1)$. AR1 provides a theoretically grounded and widely accepted proxy for quantifying ecosystem resilience. Overall, the application of CSD theory to the study of forest resilience is feasible and supported by existing theoretical and practical research[2,6,10,12,25].

Action Taken. Considering that some readers may not be familiar with the CSD concept and indicators of resilience, we improve the description of resilience (Lines 66~84 for Introduction section, we re-write this part) and add a section about the relationship between CSD theory and forest resilience indicators (Lines 258~326 in supplementary materials, i.e., Supplementary Note 8).

In addition, we used 1 km NDVI data (16-day interval, MOD13A2) and crowdsourced points to derive empirical recovery rates, which are methodologically independent from the CSD indicators, thereby providing an external benchmark to validate the CSD framework applied in this study. The methods and code provided in Smith et al (2022)[11].

The specific steps for calculating the empirical recovery rate are as follows: (1) Based on the detrended and deseasonalized KNDVI residual time series, we applied a 9-month moving window to the residuals and calculated the mean difference between the first and second halves of each window. (2) We then smoothed the resulting time series using the Savitzky-Golay filter to remove high-frequency noise. (3) Next, we identified values above the 99th percentile and marked subsequent time periods as disturbance events. For each detected disturbance event, we used a 4-month constraint to locate the local minimum in the residual time series as the recovery starting point. (4) Finally, we fitted an exponential function to the time series for four years after the starting point to calculate the exponential rate, which we regarded as the empirical recovery rate.

We compared the spatial patterns of empirical recovery rate and AR1 and found them to be generally similar (Fig. R1c, d). However, we found that in northern forests, there are discrepancies between AR1 and the empirical recovery rate (Fig. R1c, d). We would like to emphasize that the empirical recovery rate fundamentally reflects the recovery process following strong disturbances, while AR1 serves as an indicator of ecological resilience that incorporates responses to both strong and weak disturbances. Therefore, such differences are reasonable. The overall spatial patterns of the two metrics remain largely consistent. Then, we applied the KNN (N=10) approach to pair each PF point with its nearest 10 NF points and calculated the differences of empirical recovery rate between them (and AR1 difference between them). We observed that both $\Delta AR1_{(PF-NF)}$ and ΔEmpirical recovery rate$_{(PF-NF)}$ are negatively correlated with P/PET (Fig. R1e, f), consistent with the conclusions in the main text (Fig. 3).



(a)
VOD
Raw VOD
Deseasoned and detrended VOD
Residual VOD
Residual subset
Exp. fit exponent = −0.342 R^2 = 0.322
Exp. fit exponent = −0.407 R^2 = 0.435
Date
(b)
AC1
Binned medians
Binned means
Exponential fit (R^2 > 0.3)
e^r
Recovery rate
Variance
Mean theoretical variance
Power-law fit (R^2 > 0.3)
$-\sigma_r^2/2r$
Binned medians (R^2 > 0.1)
Binned medians (R^2 > 0.2)
Binned medians (R^2 > 0.3)
Binned medians (R^2 > 0.4)
Binned medians (R^2 > 0.5)
(c)
AR1-KNDVI
(d)
Emperical recovery rate
Deciles (D1~D10)
10th
30th
50th
70th
90th
(e)
y = 0.0669 – 0.0629 x
k = 10
ΔEmperical recovery rate(PF-NF)
P/PET
(f)
y = 0.0452 – 0.0244 x
k = 10
ΔAR1-KNDVI(PF-NF)
P/PET

Figure R1. The non-linear relationship between empirical recovery rate and early warning indicators (e.g., AR1). (a) Quantification of empirical recovery rates based on disturbance-recovery events in VOD. (b) Nonlinear relationship between empirical recovery rate and early warning indicators (AR1 and variance). Multiple binning and fitting approaches (medians, means, exponential and power-law fits) confirm this strong nonlinear relationship. Shaded regions indicate uncertainty or confidence intervals. Figure R1 (a-b) is modified from refs [11]. The spatial patterns of (c) AR1-KNDVI and (d) empirical recovery rate in crowding-sourcing points. (e) The AR1-KNDVI and (f) empirical recovery rate difference between PF and NF by K-Nearest Neighbors method (K = 10, similar as Figure S33, but for empirical recovery rate).

First, what are the tipping points we are considering? Is it "collapse" of the forest? how is that defined specifically for PF and NF? Defining the tipping point seems necessary to me so as to say what signal we should be using to monitor the disturbances for establishing the characterization of resilience.

R: Thanks for your suggestion. Here we will further interpret "the relation between indicators and tipping point based on CSD", including the definition of tipping point, threshold, and early-warning indicators.

A *tipping point* is a critical threshold in the environment[6,10,29]. Once that threshold is crossed, the whole system can shift (i.e., often irreversibly[2,6]) into a different equilibrium state. Therefore, a tipping point is *not* necessarily a total "collapse". It is any transition from one stable state to another[2]. According to Figure R2, Figure R2 (a) displays a linear system. As the environment changes, the system moves smoothly up or down along a trajectory. The shift is reversible. Figure R2 (b) is a non-linear system. Small environmental changes can change the system from one state to another. But this change is still reversible in principle. Figure R2 (c) shows a non-linear system with explicit thresholds. When the environment reaches threshold F_2, the system suddenly jumps to the other stable state. When the environment further changes, the system no longer changes along the curve where F_2 is located, but along a new curve (the curve F_1 changes). **That jump marks the tipping point. Such a "jump" is often irreversible.** In fact, many of the dynamic systems are those described in Figure R2 (c) (e.g., most systems are nonlinear systems[30]).

In the context of our study, which includes both planted forests (PF) and natural forests (NF), a tipping point is therefore the moment when the forest ecosystem leaves its original state and settles into a new one under strong and weak disturbances (now limited to specific disturbance)[10,31]. A familiar extreme example would be severe forest degradation by continuous and frequent fires or droughts, where a closed-canopy NF turns into shrubland (or even to grasslands). Importantly, PF and NF differ in their thresholds and responses to disturbances. PF generally have lower tipping points, meaning they are more prone to shift from their original state to secondary forests or shrublands under continuous or repeated disturbances. In contrast, NF tend to have higher tipping points and are more capable of maintaining a closed-canopy state under similar disturbance regimes. This difference mainly arises because forest management has profoundly altered the structure, composition, and soil properties of PF compared to NF[32].

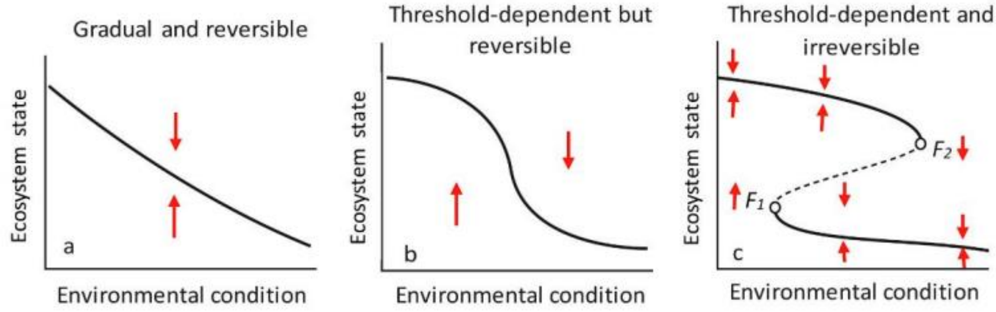


Figure R2. System changes in response to a range of conditions. (a) The system exhibits smooth, continuous changes that can be reversed as environmental conditions shift. (b) The system demonstrates abrupt transitions at certain thresholds, but these shifts can still be undone if conditions return. (c) The system undergoes abrupt, threshold-driven transitions that are not reversible—reflecting a catastrophic regime shift. Figure is from refs [10,20].

We further clarify the relationship between the process leading to tipping points and early warning indicators. When the system reaches the disturbance threshold (i.e., the tipping point), it will shift from one stable state to another, which can be represented by two attraction basins (from one attraction basin to another attraction basin)[6,25,29,33]. The deeper the attraction basin is, the higher the rate at which the system recovers from the disturbance (higher resilience or recovery rate) (Fig. R3). Therefore, *State-1* exhibits higher resilience and a stronger recovery rate than *State-2*. In the time series, *State-1* always fluctuates rapidly around the average value during disturbances. This is because the deeper attraction basin makes its recovery faster (Fig. R3b, d). In contrast, due to the shallower attraction basin, *State-2's* recovery is relatively slow, resulting in a significant deviation from the average value. However, it is eventually able to recover to the average state. These phenomena are reflected by some indicators such as autocorrelation and variance, with higher AR1 in *State-1* than in *State-2* (Fig. R3c, e).

When transitioning from *State-1* to *State-2*, the system will also approach the tipping point (Fig. R3a). Deep attraction basins will also gradually become smaller and flatter (so the small ball (i.e., system) can get closer to the tipping point under even weak disturbance)[29]. Thus, a smaller and flatter basin corresponds to a slower recovery of the system[2,6,34]. After such slow down, the system shows a longer memory, and its feature in indicators is a larger variance and correlation (e.g., AR1) between subsequent states[2,34]. Therefore, an increase in AR1 often indicates a weakening of resilience.

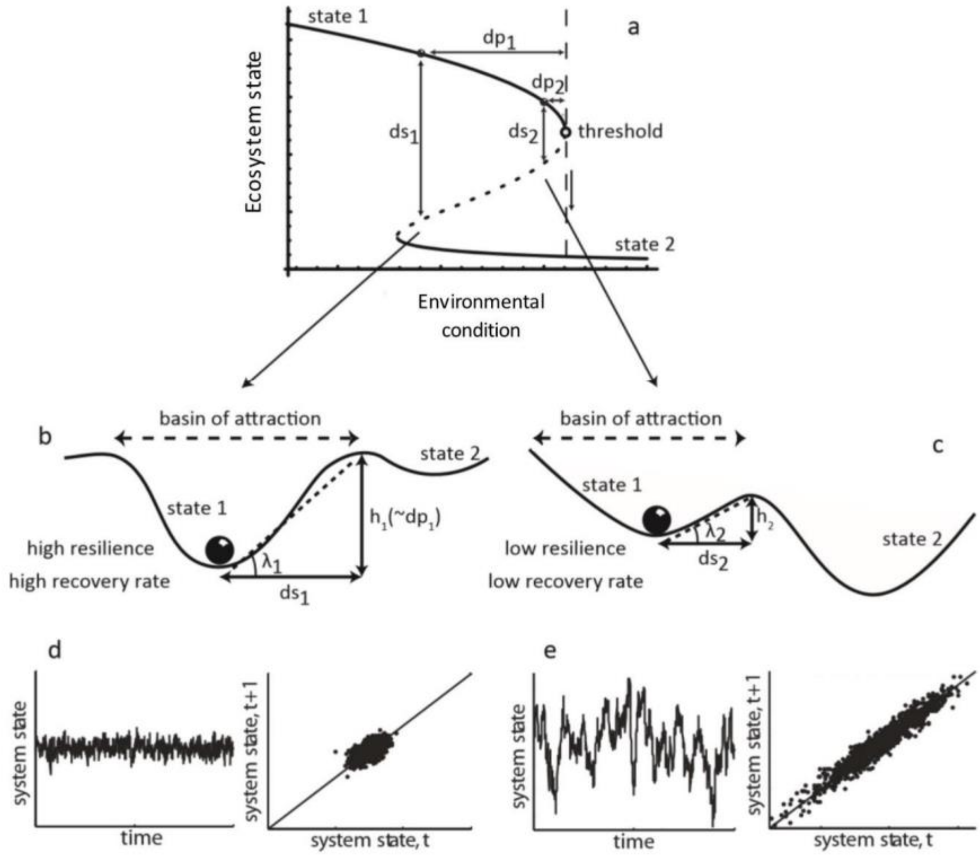


Figure R3. Conceptual illustration of ecological resilience, tipping point, and early warning indicators. (a) Ecosystem state as a function of environmental conditions, illustrating how systems can shift between alternative stable states. As the environment changes, the system may experience a gradual shift (solid line), or an abrupt, threshold-driven transition (dashed line) from *State-1* to *State-2* when a critical threshold is crossed (tipping point). (b, c) "basins of attraction" represent system resilience: A deep, steep basin (*State-1*) corresponds to high resilience and rapid recovery from disturbances. (d, e) Corresponding time series and lag plots for *State-1* (d) and *State-2* (e). In *State-1*, rapid recovery leads to small fluctuations and weak autocorrelation (low AR1, low variance). In *State-2*, slower recovery results in larger, more persistent deviations and higher autocorrelation (high AR1, high variance). Figure is from refs [7,10,20] (in Nature and Science journal).

<u>**Action taken**</u>. According to your suggestions and comments, we added supplementary note 8 and Figure S35~S36 to introduce the relationship among early-warming indicators, tipping point, and resilience. In addition, we have introduced two examples of forest management to help illustrate the above relationship. Here are the texts in the supplementary materials (Lines 258~326 in supplementary materials, i.e., Supplementary Note 8):

"**Supplementary Note 8: The relation between early-warming indicators, tipping point, and resilience.**

There are two main definitions of resilience in ecological research[30]. Engineering resilience refers to the ability of a system to recover quickly after specific and intense disturbances, such as wildfire, drought, or human activities like logging[31]. It focuses on the recovery process following a single type of strong disturbance. In contrast, ecological resilience is the ability of a system to persist and maintain its original state in the face of a continuous sequence of both strong and weak disturbances[32]. This study adopts the ecological resilience perspective, as it better reflects the continuous and various natural and human pressures faced by managed forests.

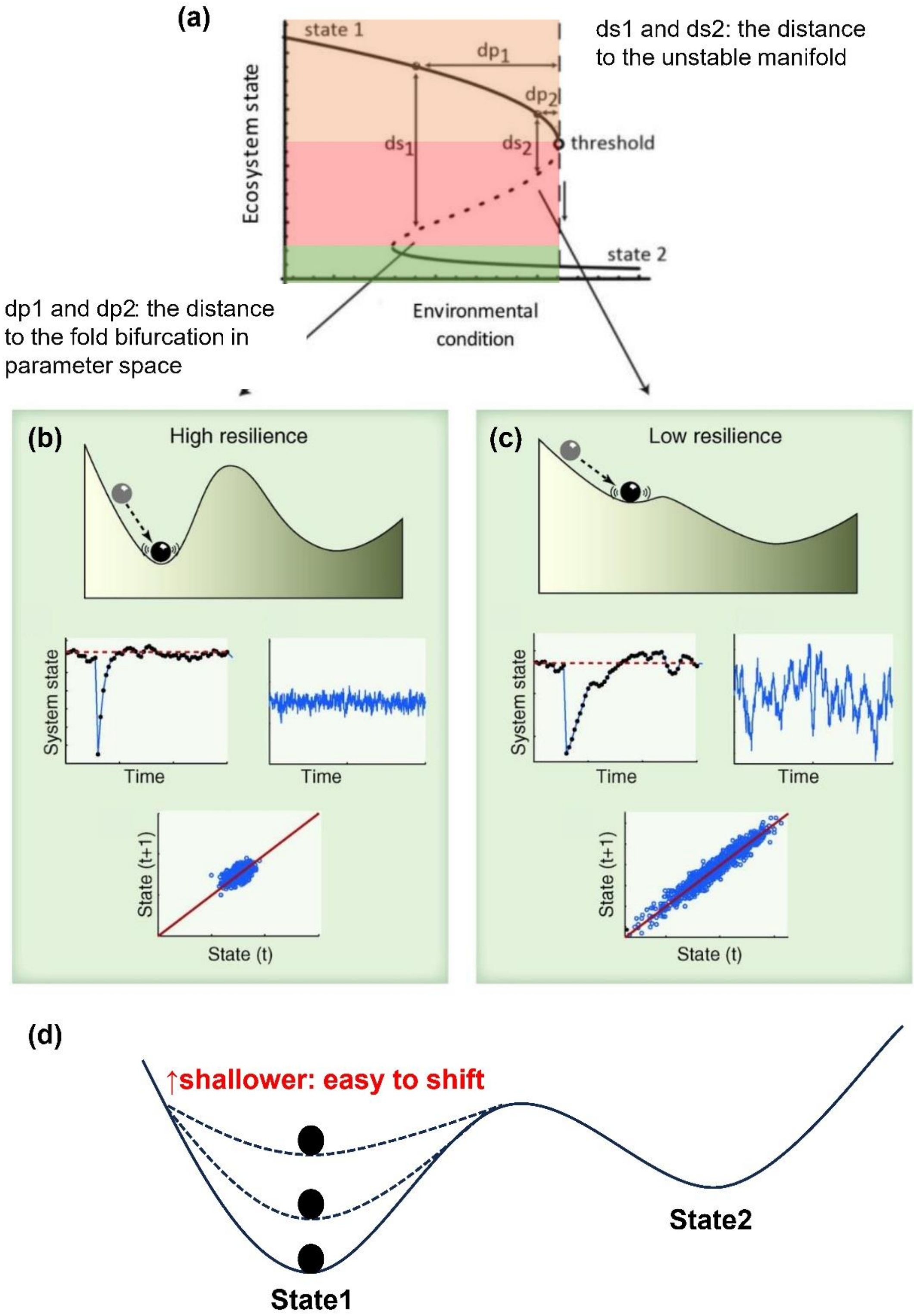
We quantify ecological resilience using indicators such as AR1, based on multiple lines of evidence. (1) Empirical relationship**:** Previous studies use the vegetation datasets to calculate the empirical recovery rate[33], which is estimated by the recovery curve of vegetation after strong disturbance. They further find that there is a strong nonlinear relationship between AR1 (variance) and the empirical recovery rate[22]. Thus, these indicators are representative of ecological resilience. (2) Mathematical foundation**:** A clear link exists between autocorrelation and the recovery rate when the recovery process is modeled as an Ornstein–Uhlenbeck (OU) process[34–36]. This mathematical framework shows that, under discrete sampling, the AR1 can be used to directly infer the system's recovery rate after disturbances ($\text{recovery rate} = -\ln(AR1)$). (3) Physical interpretation: According to critical slowing down theory, ecosystems with low resilience recover more slowly from disturbances, which is reflected in higher AR1 and variance. In such cases, disturbances leave longer memory traces in the system, indicating incomplete recovery and reduced capacity to absorb shocks. By contrast, in resilient ecosystems, disturbance effects dissipate quickly, resulting in low AR1 and variance.

These indicators are also known as early warning indicators because, according to CSD theory, they provide signals of impending irreversible shifts in a system's state. Many of the real-world systems are non-linear systems with explicit thresholds[30,37]. They often experience irreversible shifts once these thresholds are crossed. Specifically, at the beginning, when the environment changes, the system changes according to the *State-1* curve. When the environment condition reaches the threshold, the system suddenly jumps to the other stable state (from *State-1* to *State-2*). When the environmental conditions further change, the system no longer changes along the *State-1* curve, but along a new curve (the *State-2* curve) (Supplementary Fig. S35a). That jump marks the tipping point. Such a jump is often irreversible. In the context of our study, a tipping point is therefore the moment when the forest ecosystem leaves its original state and settles into a new one under disturbances (now limited to specific disturbances)[38]. An example would be severe forest degradation caused by continuous and frequent fires or droughts, where a closed-canopy forest turns into shrubland (or even to grasslands)[39]. In this regard, tipping points mark the two stable states for such example: closed-canopy forest and shrubland.

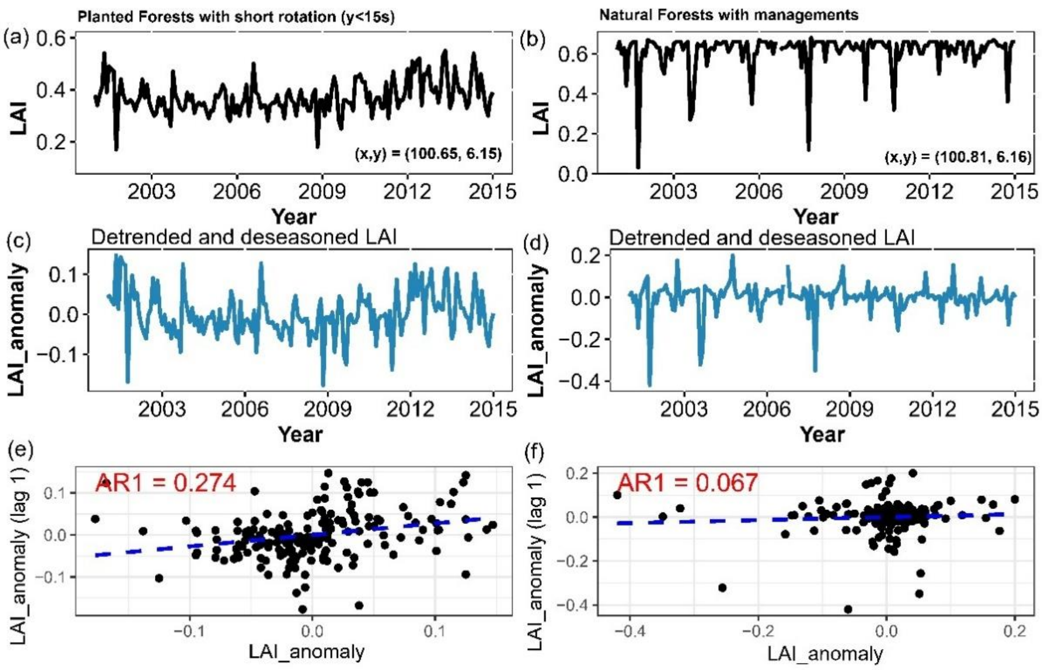
Then, we use the concept of attraction basins to illustrate how shifts between stable states are related to early warning indicators. Each state of the system is represented by a basin, and transitions between stable states are depicted as the ball moving from one basin to another (when crossing the tipping point)[37]. The depth and steepness of the attraction basin indicate the system's resilience: a deeper and steeper basin means the system can recover quickly from perturbations, reflecting higher resilience and faster recovery rates. For example, in *State-1*, where the basin is deep, the system (the ball) returns rapidly to its equilibrium position after a disturbance, resulting in small and short-lived fluctuations around the mean state. This indicates high resilience and is evident in the time series as rapid recovery with low autocorrelation and variance (Supplementary Fig. S35b). In contrast, in *State-2*, the attraction basin becomes shallower and flatter. Here, the system's recovery from disturbances is slower, leading to larger deviations from the average state and more prolonged fluctuations (Supplementary Fig. S35c). Although the system may eventually return to equilibrium, the process takes much longer, indicating low resilience. In the time series, this is reflected as higher autocorrelation (AR1) and variance, signaling that the system retains the effects of past disturbances for a longer period (“longer memory”)[40].

As the system moves from *State-1* toward *State-2* and approaches the tipping point, the basin of *State-1* continues to become shallower and flatter (Supplementary Fig. S35d). This makes it easier for even minor disturbances to push the system out of its current state, as its sensitivity to perturbations increases[37,40,41]. The progression from a deep to a shallow basin is characterized by a gradual slowdown in recovery, increasing variance and autocorrelation in the system’s dynamics (i.e., critical slowing down). These changes serve as early warning signals for reduced resilience and impending regime shifts.

Our analysis of LAI anomaly (after detrending and deseasonalization) in two types of forests provides support for the theoretical concepts of resilience and critical slowing down. Specifically, planted forests with short rotation periods (years < 15) exhibit higher AR1 values than nearby managed natural forests (Supplementary Fig. S36). This relatively higher AR1 value suggests that the LAI of planted forests retains the effects of past disturbances for a longer period—indicating a “longer memory” and slower recovery rate (Supplementary Fig. S36 c, e). In contrast, LAI anomalies in managed natural forests tend to dissipate more rapidly, implying a shorter memory and higher ecological resilience (Supplementary Fig. S36 d, f). These observations align with the CSD theory, suggesting that the rotational management practices in planted forests may reduce ecological resilience.”



Supplementary Figure S35. Conceptual illustration of resilience, early warning indicators, and tipping points in ecological systems. (a) Ecosystem state trajectories under changing environmental conditions. This panel depicts how a system can transition from one stable state (*State-1*) to another (*State-2*) when a critical environmental threshold (tipping point) is crossed. ds_x is the distance to the unstable manifold (dotted line in (a)), representing the maximum disturbance the system can withstand before shifting states. dp_x is the distance to the fold bifurcation in parameter space (i.e., environmental conditions). (b, c) Basins of attraction illustrating system resilience and corresponding time series properties: (b) High resilience: a deep basin reflects strong recovery capacity, characterized by small fluctuations, low autocorrelation (AR1), and low variance. (c) Low resilience: a shallower basin indicates slower recovery, with larger fluctuations, higher AR1, and increased variance. (d) Progressive shallowing of the basin of attraction: as the system approaches a tipping point, the basin becomes shallower and flatter, making the system recover more slowly after disturbances and signaling declining resilience. These changes in time series properties, particularly increases in AR1 and variance, serve as early warning indicators of impending critical transitions and regime shifts in ecosystems. Figure is modified from refs [11,30,42].



Supplementary Figure S36. Comparison of LAI anomaly time series and AR1 for nearby different forest management types. (a, b) LAI time series for a planted forest with short-rotation management (100.65°E, 6.15°N) and a nearby natural forest with management (100.81°E, 6.16°N). (c, d) Corresponding LAI anomaly after detrending and deseasonalization. (e, f) Autoregressive lag-1 showing stronger temporal memory in the short-rotation planted forest (AR1-LAI = 0.274) compared to the managed natural forest (AR1-LAI = 0.067)

Then, the next question is: what specifically are we considering as "disturbances"? Windthrows? Fires? bark beetle infestations? clear-cuts? Later in the discussion it states (L500) "abrupt events and human disturbances lead to non-stationary conditions" and that this limits the accuracy. So I do not exactly understand what disturbances we are trying to see in the signal.

R: Thanks for your valuable comments and suggestions. **The definition of specific and explicit strong disturbances corresponds to the concept of engineering resilience, whereas this study adopts the concept of ecological resilience.** According to the definition of ecological resilience[21,25], disturbances encompass a continuum of events with varying intensities, regardless of their specific types. "Disturbances" refer to the broad continuum of events that can influence forest ecosystem states, including both natural (e.g., droughts, windthrows, fires, pest outbreaks) and anthropogenic (e.g., clear-cuts, afforestation, restoration, land-use change) drivers. Our approach quantifies resilience as the system's capacity to maintain its state and stability under a continuous sequence of natural and anthropogenic disturbances, rather than recovery from any single specific disturbance. In fact, managed forests are exposed to a continuous sequence of anthropogenic and natural disturbances rather than a single specific disturbance. In addition, this cumulative approach is a core principle of ecological resilience and is well-suited for global-scale monitoring using satellite remote sensing data. While this means we cannot attribute observed changes to any one specific event, it enables us to capture the overall recovery capacity of the system under real-world conditions, as recommended by the CSD framework.

Drought events are typically identified by specific thresholds (e.g., PDSI < -2), but even moderate water deficits (-2 < PDSI < 0) can adversely affect vegetation in some regions, and the system's recovery from these less severe but more frequent disturbances is also relevant for quantifying resilience. This highlights that not only major disturbances such as windthrows or fires, but also smaller and more frequent events, can influence ecosystem stability and resilience. The method in this study is to quantify resilience as the capacity of the system to recover from a continuous sequence of disturbances of different types and intensities consistent with the ecological resilience concept and the CSD framework.

In addition, we also calculated the empirical response rate and CSD resilience indicators based on crowdsourced data to verify the robustness of our results. The results show the strong relationship between the empirical recovery rate and the CSD indicator (Fig. 1Rc-f).

We further clarify the meaning of the statement, "abrupt events and human disturbances lead to non-stationary conditions". Under intensifying climate change, extreme events can become frequent. Frequent extreme events continuously damage the vegetation. Even after we remove seasonality and long-term trends, the residual series may not meet the stationarity assumption that CSD indicators rely on[35]. When this assumption is not satisfied, early-warning metrics such as AR1 and variance may yield unreliable signals of resilience. For this reason, the application of the CSD framework to future projections (e.g., CMIP6 scenarios) requires particular caution. Recent studies have shown that resilience indicators derived from CMIP6 models exhibit substantial uncertainties and do not consistently reveal coherent trends[35,36].

Despite these challenges, the application of CSD-based indicators to current remote sensing vegetation observations remains valid and widely accepted. Previous studies have successfully employed this approach using satellite remote sensing data[10,12,15,31], demonstrating that the stationarity assumption is sufficiently met to extract meaningful early-warning indicators.

Action taken. According to your suggestions and comments, we have deleted the related sentence (that could mislead the readers).

This adds questions then to whether the signal used are appropriate in terms of (1) what they measure, (2) over what spatial supports, and (3) with what temporal frequency.

For point (1), all metrics used fundamentally relate to temporal variations in greenness (whether they be LAI, kNDVI, fAPAR, GPP, they are reflect mainly greenness and the information comes mostly from NIR and Red reflectances).

R: We agree that the choice of vegetation index can influence the assessment of ecosystem resilience. LAI, kNDVI, fAPAR, and GPP represent complementary aspects of vegetation structure and function, and are widely used in monitoring vegetation dynamics and recovery processes in response to climate and human disturbances, such as drought and fire (engineering resilience)[24]. Many recent studies have successfully applied CSD-based early warning indicators to these optical indices (e.g., LAI, GPP, kNDVI) to estimate ecological resilience[13,19].

(1) The coarse spatial resolution of microwave-based indices. Microwave-based indices, such as vegetation optical depth (VOD), offer the advantage of penetrating the canopy and capturing total biomass, but are limited by their coarse spatial resolution (typically ~25 km)[37]. Such low resolution would impair our ability to compare resilience among different forest management types at the landscape scale. For this reason, VOD was not used in our main analysis. Instead, we prioritized optical indices, which provide much higher spatial resolution and are more suitable for investigating forest management effects across large spatial extents.

(2) The discussion about the reliable assessment of global forest resilience. Although a recent study[15] shows that when using the optical index to calculate the resilience of vegetation in tropical areas (i.e., issues like optical vegetation index saturation), the deviation between λAR1 and λVar is relatively large, suggesting these indicators should be interpreted with caution in such contexts. **Our study includes indices that are not easily saturated, such as kNDVI, which showed similar patterns with other indices (Fig. 2).** Our results indicate that the paired grids of PF and NF are mainly located in the non-tropical forest areas of the Northern Hemisphere (Fig. 2 and Fig. S7, the latitude of paired grids is mostly over 25°N, not in the tropical areas). In addition, we also used the process-based model output GPP_PML (based on Penman–Monteith–Leuning model), which exhibited similar patterns (Fig. 2).

Furthermore, we used multiple early warning indices (AR1, λAR1, and λVar) and obtained similar results, which proved the robustness of our results (all show similar patterns by various early-warning indicators) (Figs. 2). Despite inherent differences among these indices, our global resilience maps based on optical indices (Fig. S2) exhibit spatial patterns that are highly consistent with those derived from VOD in previous studies[11]. Other related studies have also calculated global vegetation resilience maps[10,15,19] through different vegetation indices, all showing similar patterns with ours (Fig. R6). These results all indicate that using the greenness index in this study can capture the global resilience.

Action taken. We mentioned this limitation in the discussion section, as following (Lines 491~500):

"We recognize that spectral greenness metrics employed here, such as LAI, are prone to saturation in dense forests, an effect that can potentially bias resilience estimation[35]. Alternatively, vegetation indices based on microwave sensors, such as Vegetation Optical Depth (VOD), can be valuable alternatives; however, their coarse spatial resolution (e.g., 0.25°) limits their ability to detect fine-scale management-related patterns. To test the robustness of our findings with respect to possible saturation effects, we employ kNDVI, a recently developed spectral vegetation index that has been documented to be more closely related to primary productivity and resistant to saturation compared to similar products based on surface reflectance of vegetation[51]."

Are changes in greenness over 5 by 5km (point (2)) detectable with a frequency of one or twice a month (point (3)) really capturing the target "disturbances" that are not man-made and not abrupt? Is this the approapriate tool for this question?

R: Here, we further respond to the reviewer's comment by addressing separately the two questions.

Question 1: Is the signal used appropriate in terms of temporal resolution? Are using greenness datasets with a frequency of one or twice a month really capturing the target "disturbances"?

(1) **Regarding the temporal resolution, we have conducted relevant sensitivity tests (Fig. S22~S25).** We calculate the AR1, λAR1, and λVar at different temporal resolution, showing similar patterns. For example, we tested the monthly LAI data synthesized by MVC and the LAI data based on the sampling scale of the original data (8 days) respectively, and calculated the relevant early warning indicators. Then, by comparing the differences in resilience under different forest management conditions, similar patterns were discovered. It indicates the temporal resolution may have little impact on the resilience estimation.

(2) **Early warning indicators are reliable for detecting changes in forest resilience and disturbance events using remote sensing time series.** In addition, as demonstrated in our example figure, after removing seasonality and trends from the LAI time series (i.e., monthly data), we identified two distinct disturbance-induced declines in LAI. In both periods, the early warning indicators (such as AR1 and λAR1) showed significant increases prior to the disturbances (Fig. R4). This illustrates that CSD-based early warning indicators derived from observational vegetation indices (monthly or bi-weekly datasets) can effectively capture disturbance-induced changes in ecosystem resilience. Similar conclusions have been reported in a previous study[15]. For example, an existing study used NDVI data with a 16-day temporal resolution, applied deseasoning and detrending, and calculated the corresponding AR1 index[15]. The deseasoned and detrended NDVI time series exhibited larger deviations from the mean (i.e., disturbances) during 2012–2016, resulting in a rapid increase in AR1 over this period (Fig. R5).

(3) **Method is based on previous studies, and the global patterns of resilience are similar to those of previous studies.** For long-term and large-scale forest remote sensing, sampling frequencies ranging from bi-weekly to monthly scales are more appropriate for CSD framework. Previous studies also used the same datasets (temporal resolution from bi-weekly to monthly) and processing methods to calculated the early-warning indicators[10,11,14,31]. The method of this study is based on these studies. We estimated the long-term average resilience based on the entire observation (15 years). The global resilience pattern calculated based on the method in this study (Fig. S2) is consistent with previous studies[10,11,15,19] (Fig. R6). This consistency further validates the reliability of our methodological framework and results.

In summary, we conclude that bi-weekly to monthly sampling intervals are sufficient and effective for capturing ecosystem resilience dynamics using CSD-based indicators. More details and additional sensitivity results can be found in the following responses (see the specific questions).

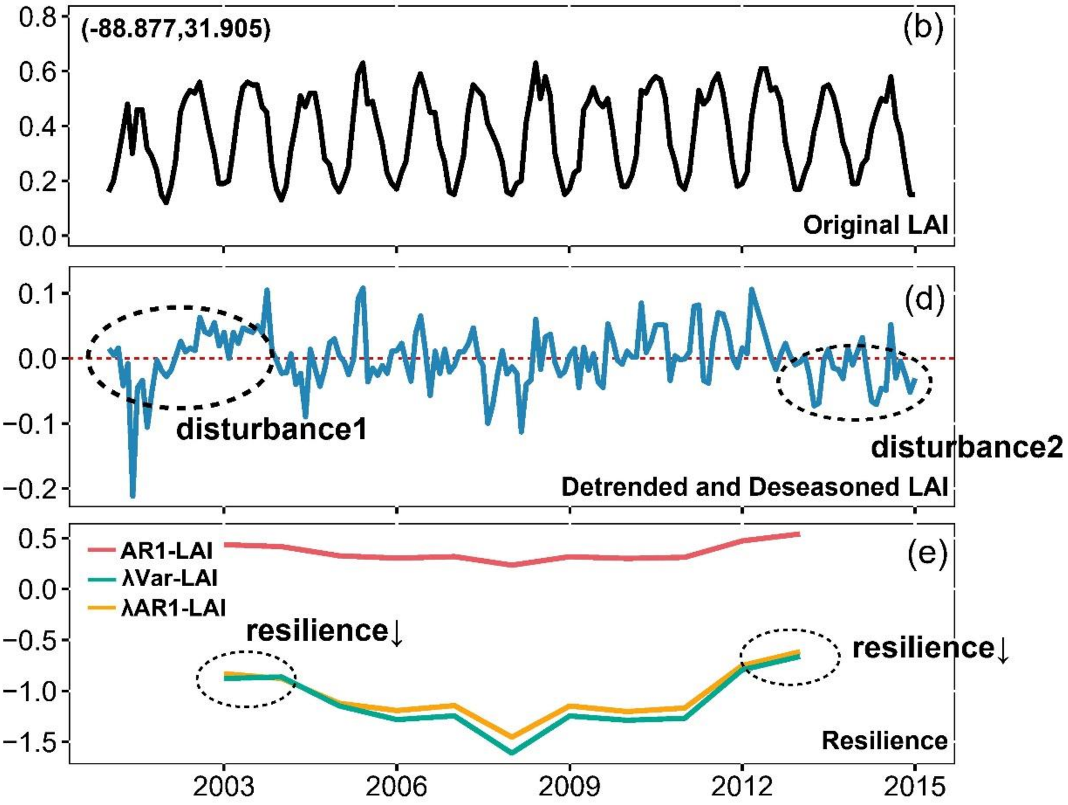


Figure R4. Figure from Figure1 in the main text. (b) Original LAI time series at the given location, showing strong seasonal patterns. (d) Detrended and deseasonalized LAI, with two major disturbance events (disturbance1 and disturbance2) indicated by sudden declines. (e) Early warning indicators (AR1-LAI, λVar-LAI, and λAR1-LAI) calculated from the processed LAI time series. Notably, all indicators show significant increases (i.e., resilience decline) preceding and during the disturbance periods, demonstrating the sensitivity of CSD-based metrics to disturbance-induced changes in ecosystem resilience.

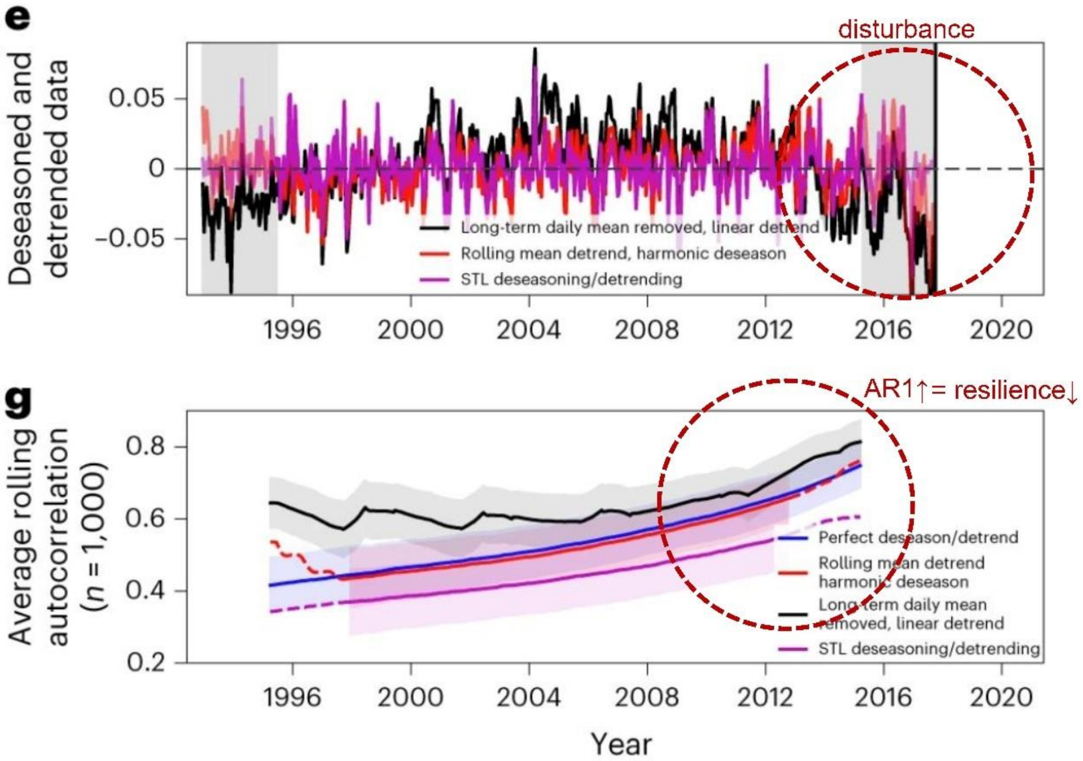


Figure R5. Detrending and deseasoning of NDVI time series and corresponding rolling AR1 calculations. (e) NDVI time series (16-day resolution) after removing seasonality and trends using different methods: long-term daily mean removal with linear detrending (black), rolling mean detrending with harmonic deseasoning (red), and STL-based deseasoning/detrending (magenta).
(g) Average rolling autocorrelation (AR1) calculated over a moving window (n = 1,000) for each preprocessing method. The comparison shows that different deseasoning and detrending approaches can impact the estimated AR1 values, highlighting the importance of preprocessing choices when applying early warning indicators to remote sensing data. This figure is modified from ref[15].

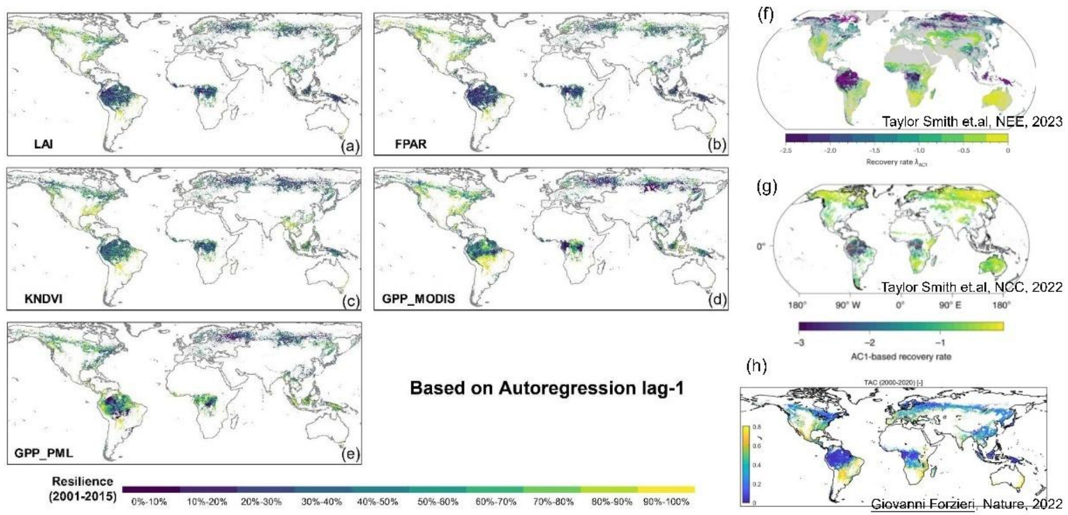


Figure R6. Comparison resilience estimates between this study and previous studies. (a-e) The global spatial distribution of AR1-LAI estimated by this study's method. (f) The resilience distribution of λAR1-kNDVI estimated by Smith et.al[15] (published in Nature Evolution and Ecology journal). (g) The resilience distribution of AR1-VOD estimated by Smith et.al[11] (published in Nature Climate Change journal). (h) The resilience distribution of AR1-kNDVI estimated by Forzieri et.al[10] (published in Nature journal).

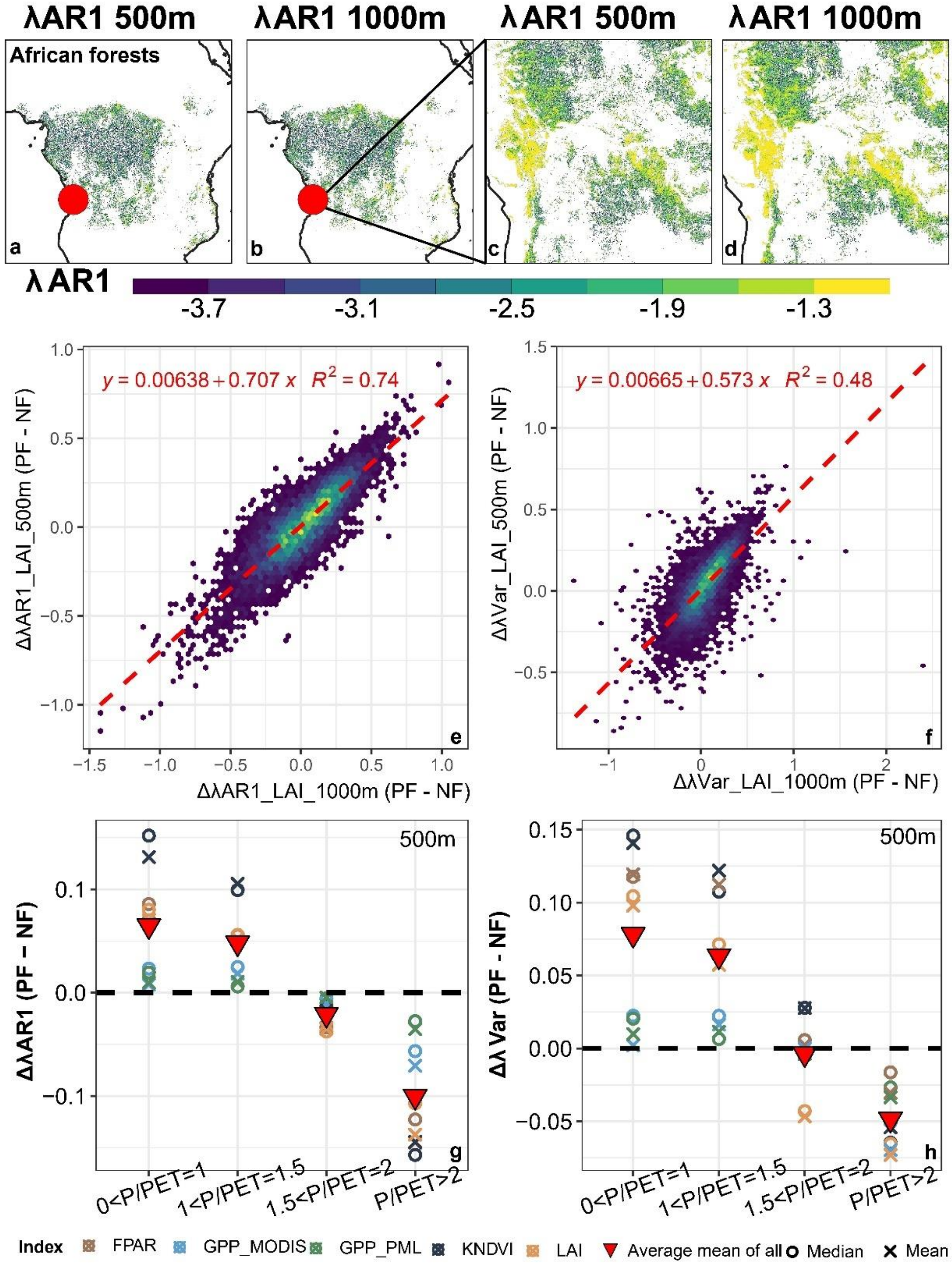
Question 2: Is the signal used appropriate in terms of spatial resolution? Are changes in greenness over 5 by 5km (point (2)) detectable with a frequency?

(1) **First, we would like to clarify that we have used 1km resolution data in the main text, instead of 5 km.**

Action taken. Regarding spatial resolution, we conducted a sensitivity test of our result: 500 m and 1 km (Fig. S29). For example, we calculated early warning indicators using both the 500 m MVC-synthesized LAI data and the 1 km MVC-synthesized LAI data, and found similar spatial patterns in the results (Fig. S29). Similarly, we resampled the GFMD and GHFC maps to 500 m resolution and performed moving-window analyses, generating the resilience between PF and NF. The patterns by 500m datasets consistent with those from the 1 km analysis presented in the main text (Fig. S29). These sensitivity tests confirm that our main findings are robust to the choice of spatial resolution. The related discussion and method are displayed in the supplementary materials (Supplementary Note 5).

(2) **The reasons of using 500 m or 1 km datasets in this study.** Given the global scale of this comparative study, we used spatial resolutions of 500 m and 1 km. Continuing to improve the spatial resolution will significantly increase our computing cost and time. Moreover, apart from MODIS-based and Landsat-based products, there is a lack of globally continuous high-resolution remote sensing data for the 2001–2015 period (i.e., open-access). Although Landsat data offer higher spatial resolution, their temporal frequency is only every 16 days, and for some locations—especially those affected by cloud cover—observations may be unavailable for an entire month (or even multiple month)[38]. Additionally, the scan line corrector (SLC) failure of Landsat 7 introduces further data gaps and uncertainties[39].

(3) **We also emphasize that spatial resolution and temporal resolution are mutually restrictive in remote sensing data[40,41].** It is important to note that there is an inherent trade-off between spatial and temporal resolution in remote sensing data. Higher spatial resolution satellites often have lower temporal coverage and more missing data, which can limit the frequency and continuity of earth observations. This, in turn, would negatively impact the calculation of resilience indicators that require consistent time series data[14,15]. For these reasons, our used spatial resolutions represent (i.e., the MODIS vegetation datasets) an optimal balance between spatial detail, temporal continuity, and data reliability for large-scale resilience assessment.



Supplementary Figure S29. Effect of spatial resolution on resilience estimates at 500 m and 1000 m. (a–d) Spatial distribution of λAR1-LAI in African forests at 500 m and 1000 m resolution. (e, f) Correlations of ΔλAR1-LAI $_{(PF,\ NF)}$ and ΔλVar-LAI$_{(PF,\ NF)}$ between the two resolutions. (g, h) Resilience differences (PF − NF) along the P/PET gradient at 500 m resolution.

To bring a bit more of a specific detail of why this is problematic for this question, consider that managed forest will be more subjected to clear cuts by definition. These will entail stronger drops in greenness than for natural forests, so when this is averaged of coarse spatial resolutino of 5km, the mixed signal will be more strongly affected (on average) for the managed forest than natural forest.

R: The reviewer is concerned that the above core conclusions might be more significant at coarser resolutions (spatial scale effect). However, the scales on which our research is based **are 500 m and 1 km** (Fig. S29 and Figs. 2-5). At these finer resolutions, the risk of spatial mixing between highly disturbed managed forests and undisturbed natural forests is substantially minimized. The spatial scale effect was explicitly tested, and we found that the main findings persist regardless of spatial resolution at 500 m or 1km. Thus, the pattern is not a false positive driven by spatial scale (i.e., spatial resolution), but an intrinsic property of forest management practices.

The reviewer raised concerns that differences in disturbance regimes and ecological contexts between planted forests and natural forests may affect their resilience gaps. i.e., managed forest will be more subjected to clear cuts. **The influence of land cover change on the resilience is removed by controlling forest translations (e.g., forest to non-forest).** To further mitigate large logging's effects, we conducted analyses using the GHFC and GFMD maps to identify and exclude pixels with land cover changes, including areas transitioning between forest and non-forest states from 2000 to 2015 (see method, Lines 554~563). This approach substantially mitigates the influence of land-use change and ensures continuous human management here.

Moreover, the planted forests (PF) considered in our study are defined as forests established for ecological restoration purposes, typically with rotations longer than 15 years. As a result, these PF do not experience frequent or intensive logging[42]. The resilience differences detected in our analysis are primarily attributable to differences in forest management types rather than large-scale logging or land use change.

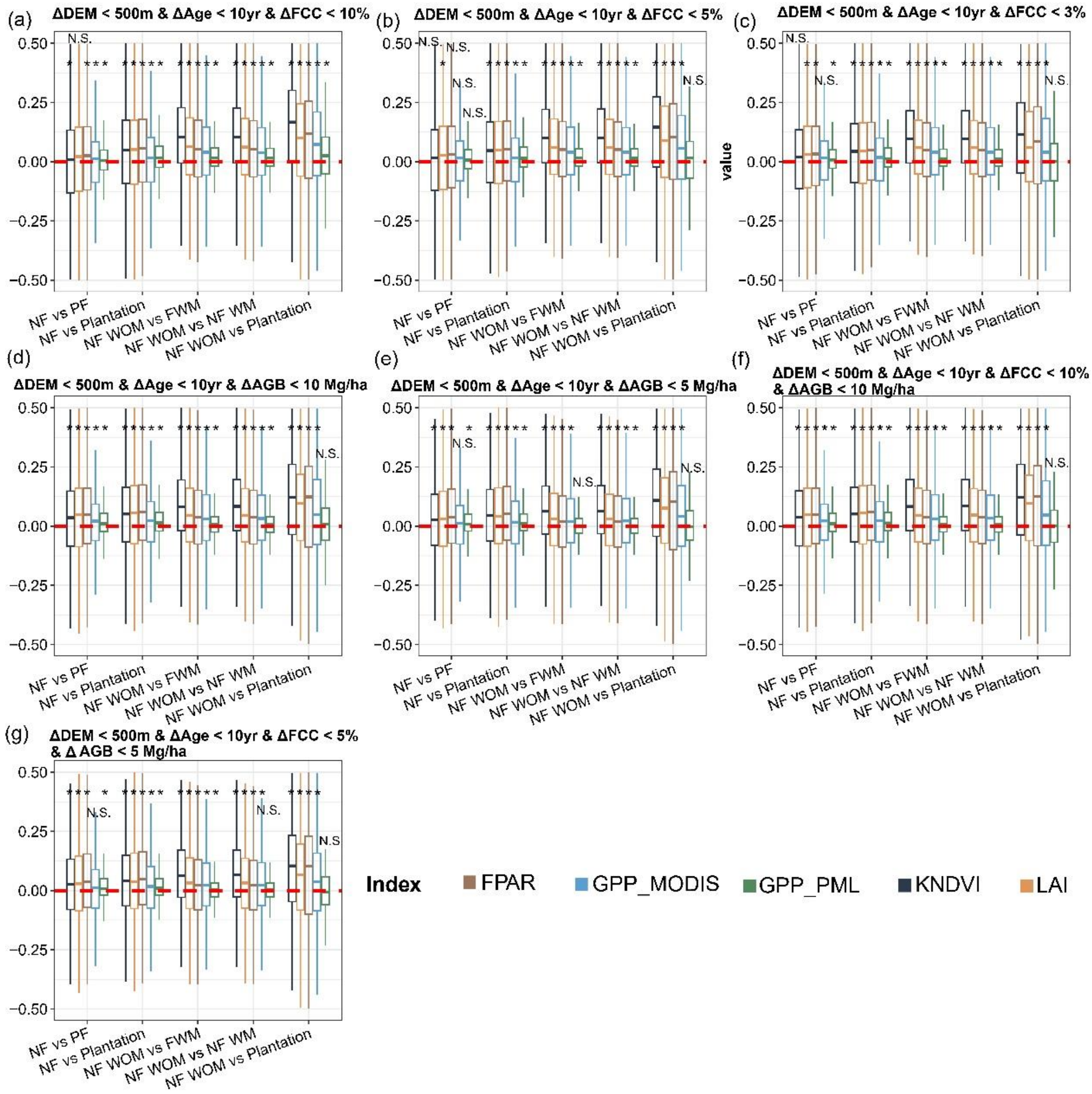
Action taken. To address this concern (the influence of logging difference between PF and NF on the resilience gap), we conducted a sensitivity experiment by introducing two additional filtering criteria within the moving window (Fig. S21). The detailed discussion and method are displayed in the supplementary materials (Supplementary Note 3). Here is the method of sensitivity test, taking PF and NF as examples.

(1) To control for disturbance like deforestation or afforestation, we calculated changes in forest cover for both planted and natural forests. The change in forest cover from 2001 to 2015 was derived from the Global Forest Cover Change (GFCC) dataset[43], which provides tree cover information at a 30-meter spatial resolution for the period 2000–2015. The overall change in forest cover occurred over the period 2000-2015 (e.g., $FCC = FC_{2015} - FC_{2000}$) is used to derive the reference disturbance regime map. We then selected only those paired grid cells where the difference in forest cover change between the two types was limited to 10%, 5%, or 3% ($|\Delta FCC_{(PF-NF)}| < 10\%, 5\%, \text{or } 3\%$).

(2) To further enhance comparability, we estimated aboveground biomass (AGB) for both planted and natural forests and selected paired grids with biomass differences of less than 10 Mg/ha or 5

Mg/ha ($|\Delta AGB_{(PF-NF)}| < 10$ Mg/ha or 5 Mg/ha). The AGB data were obtained from the NASA/ORNL biomass dataset[44], which has a spatial resolution of 300 meters.

These filtering steps ensured not only comparable climate conditions (as already controlled by the spatial window) but also similar disturbance regimes and ecological contexts. We further applied this method to comparisons involving other forest management types and found that our main conclusions remained robust. The influence of disturbance regimes and ecological differences on our core findings was limited (Fig. S21). The related texts are displayed (Lines 794~810 and Lines 68~90 in the supplementary materials).



Supplementary Figure S21. Sensitivity test of different filtering thresholds in paired grid method. Each panel shows ΔAR1 between two management types under constraints on elevation difference (ΔDEM), forest age difference (ΔAge), the difference in forest cover change (ΔFCC), and aboveground biomass difference (ΔAGB): (a) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 10%; (b) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 5%; (c) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 3%; (d)

ΔDEM < 500 m, ΔAge < 10 yr, ΔAGB < 10 Mg ha^{-1}; (e) ΔDEM < 500 m, ΔAge < 10 yr, ΔAGB < 5 Mg ha^{-1}; (f) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 10% and ΔAGB < 10 Mg ha^{-1}; (g) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 5% and ΔAGB < 5 Mg ha^{-1}. Boxplots summarize results across vegetation indices (FPAR, GPP_MODIS, GPP_PML, kNDVI, LAI). Significance levels are indicated by asterisks (*: $p < 0.05$, N.S. = not significant).

First, is any normalization used here? will the AR1 metric not be artificially affected by this?

R: We did not apply normalization. Our data processing followed previous studies: we used linear regression to remove long-term trends and calculated anomalies to remove seasonality. No further normalization was performed.

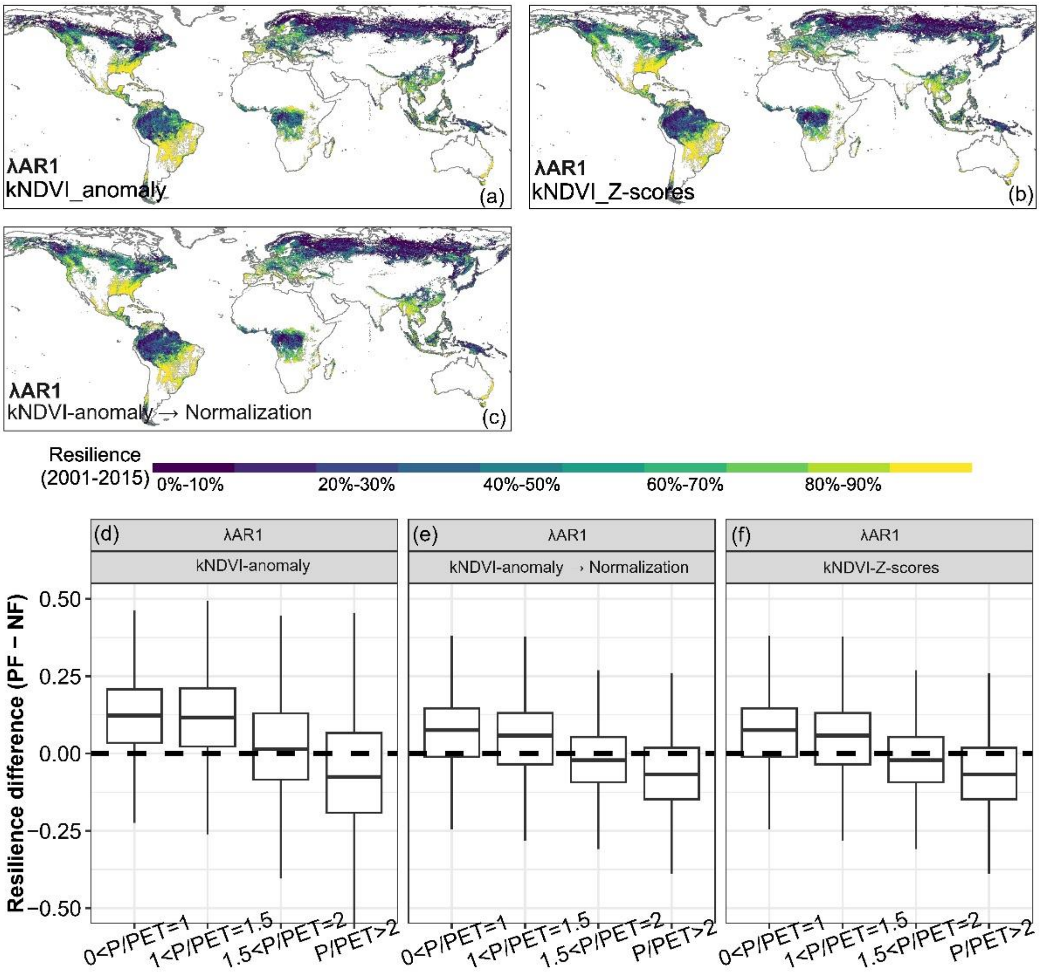
Action taken. To assess whether normalization could affect our conclusions, we conducted the following sensitivity analyses. Here we focused on its effect on λAR1. We compared λAR1 derived from three kNDVI time series: (1) an anomaly series, obtained by subtracting the multi-year mean of each month from the corresponding monthly value; (2) a standardized series, where each monthly value was expressed as its deviation from the multi-year monthly mean divided by the corresponding multi-year monthly standard deviation (Z-scores); and (3) a normalized anomaly series, where each monthly anomaly was further rescaled to the [0,1] range using the minimum and maximum anomalies of the same month. Here are the related equations:

$$kNDVI\ \ anomaly = kNDVI - kNDVI_{mean} \quad (1)$$

$$kNDVI\ \ Z - scores = \frac{kNDVI - kNDVI_{mean}}{kNDVI_{std}} \quad (2)$$

$$Normalized\ \ kNDVI\ \ anomaly = \frac{(kNDVI - kNDVI_{mean}) - (kNDVI_{min} - kNDVI_{mean})}{(kNDVI_{min} - kNDVI_{mean}) - (kNDVI_{max} - kNDVI_{mean})} \quad (3)$$

Where $kNDVI_{mean}$, $kNDVI_{std}$, $kNDVI_{min}$ and $kNDVI_{max}$ denote the multi-year monthly mean, standard deviation, minimum, and maximum, respectively, for the corresponding month. Where $kNDVI$ is the values for a specific month. Then, we applied the sliding-window algorithm to compare differences in λAR1 between PF and NF. Results show the normalization/standardization has minimal impact on the key findings (Fig. S26). The detailed discussion and method are displayed in the supplementary materials (Lines 824~830 and Lines 124~148 in the supplementary materials, i.e., Supplementary Note 4).



Supplementary Figure S26. Effect of normalization and standardization methods on resilience estimation. (a–c) Spatial distribution of λAR1 derived from kNDVI anomalies (a), kNDVI Z-scores (b), and normalized kNDVI anomalies (c) for 2001–2015. (d–f) Resilience differences between PF and NF, expressed as ΔλAR1-kNDVI $_{(PF\ -\ NF)}$, estimated with kNDVI anomalies, normalized anomalies, and Z-scores under different P/PET gradients.

Then, the recovery could be faster in a planted forest if fast-growing species are planted (and assisted, with protection from herbivory for instance) than if natural regeneration occurs. This would be reflected in the metric by suggesting managed forest would have higher resilience (as faster recovery), and yet this is counter to what the results suggest. If such clear-cuts events should not be considered (as suggested by some phrases in the text cited above), then how is this prevented from affecting the results and the conclusions?

R: **While PF may exhibit faster growth rates than NF, this does not necessarily equate to higher ecological resilience.** The long-term signals caused by greening (i.e., long-term growth trends) have been removed in the de-trend method (i.e., before calculating the resilience indicators, de-seasonalization and de-trending methods need to be performed on the data)[10]. Resilience metrics derived from CSD theory (such as AR1, λAR1, and λVar) capture the system's ability to maintain original state from perturbations, rather than the speed of greening[17].

(1) Rapid greening ≠ higher resilience. Many previous studies have demonstrated that rapid increases in vegetation greenness do not necessarily indicate enhanced ecological resilience (e.g., at the global scale[10,19], in China's Loess Plateau[17], and in the Arctic-Boreal Vulnerability Experiment region[31]). Globally, forest ecosystems are overall becoming more productive but also more fragile[10]. Therefore, it is important to note that resilience and productivity (greenness) represent two distinct dimensions.

On China's Loess Plateau, large-scale afforestation and grassland restoration have dramatically increased vegetation cover over the past two decades[45]. However, recent analyses show that this "greening" was often achieved through monocultures of exotic or fast-growing species like grass, resulting in communities with low species and functional diversity[46,47]. While these plantings can quickly cover bare soil and increase NDVI, they remain structurally simplified and thus ecologically fragile. As a result, such areas are highly vulnerable to pests, droughts, or extreme rainfall because they lack the diversity needed to buffer stress and facilitate recovery.

This pattern is not unique to the Loess Plateau. In the Arctic-Boreal Vulnerability Experiment (ABoVE) region, observed "greening" often reflects the expansion of sedge meadows or shrublands dominated by just a few pioneer species[31,48]. These simplified systems provide limited buffering against disturbances, so their apparent recovery in remote sensing metrics may mask an underlying erosion of resilience.

Such rapid greening can be considered a form of structural overshoot[49], where aboveground vegetation cover increases quickly. while supporting structures, such as root depth, soil moisture reserves, nutrient cycling, and species diversity, do not increase at the same pace. This mismatch creates an unstable state. The overshoot system appears to have high values in greenness, but its underlying resilience remains weak. Consequently, these overshoot dynamics make ecosystems highly vulnerable to subsequent droughts, pests, or extreme events.

Moreover, other studies highlight that greener vegetation is not always healthier or more stable. Even if leaf area rebounds rapidly, limited rooting depth, soil moisture depletion, or nutrient imbalances can undermine long-term stability[50]. For example, in water-limited environments, dense artificial greening may overdraw soil moisture and set up drought feedbacks, ultimately increasing the risk of vegetation die-off[46,51]. Overall, robust ecological resilience requires more than rapid green cover; it needs an ecosystem with sufficient diversity and structural complexity to absorb shocks and regenerate.

(2) As for PF and NF. PF usually rely on fast-growing tree species that shoot up and leaf out rapidly, producing a sharp NDVI increasing. The tree species are alike and diversity is low, PF lacks the different defenses to cope with threats like pests or droughts (low resilience)[32,52]. Ecological resilience rests on diversity-driven stability and a various of disturbance-mitigation pathways.

In our study, we explicitly identified and excluded all pixels that experienced transitions between forest and non-forest states during 2000–2015, as described in the methods section. The resilience metrics in our analysis are designed to quantify the system's capacity to resist and maintain its original state under continuous disturbances, including both strong and weak events. In our paired-grid, the detected resilience differences primarily reflect the effects of forest management type. For example, planted forests and managed natural forests are more frequently subjected to human interventions such as logging, which can disrupt structural complexity and, consequently, reduce resilience.

Action Taken. As the reviewer was concerned that planted forests might exhibit higher resilience simply because they grow faster, we conducted additional sensitivity analyses to control for potential differences in vegetation cover change. To further ensure a rational comparison between forest management types, we conducted **additional sensitivity analyses by restricting paired grid cells to those with similar changes in coverage and aboveground biomass (Fig. S21)**. This approach ensures that observed resilience differences are not driven by ecological background difference or unbalanced disturbance regimes. The resilience difference is driven by the management. The sensitivity-test results of these analyses were consistent with our main findings. The detailed discussion and method are displayed in the supplementary materials (Lines 69~90 in the supplementary materials, i.e., Supplementary Note 3).

I can go on with more concerns, such as the questioning of whether there are enough samples in time to have sufficient frequency to sample the stable state as suggested in Figure 2c and 2d of Scheffer et al (2015), mentioned by the authors as the reference. With monthly data over 20 years, I have some serious doubts that this can apply so seamlessly. I would think you would then need daily data from satellite, which is of course available.

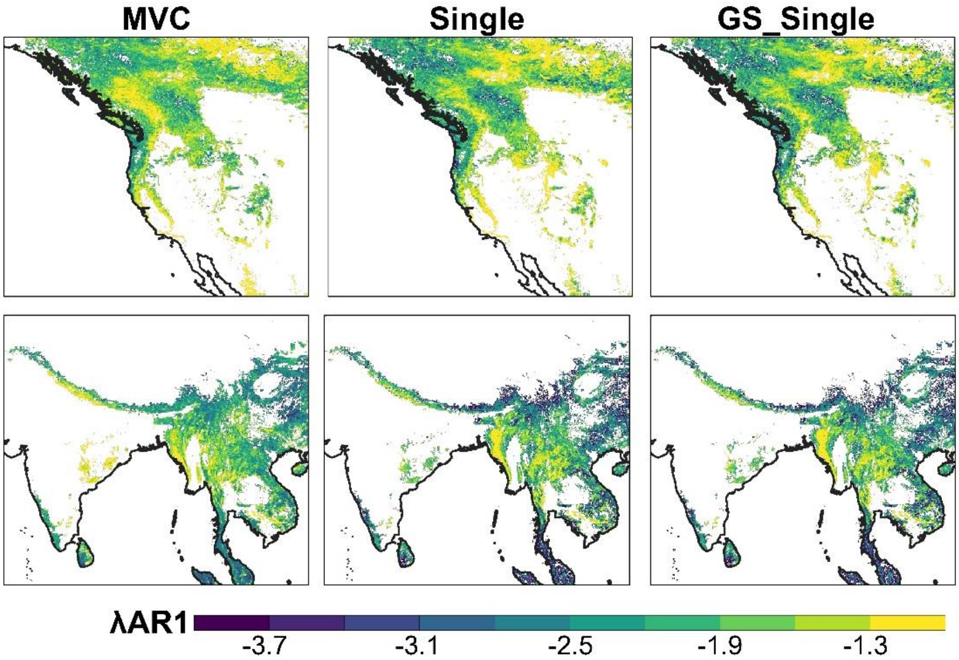
R: The reviewer raised an important concern regarding whether the choice of temporal sampling frequency may influence our results, referencing Scheffer et al.[7] as support for using daily-scale data. Below, we provide detailed responses:

(1) Regarding temporal sampling frequency, the appropriate temporal sampling frequency for early warning indicators should match the inherent dynamics of different systems. Scheffer et al.[6] employed CSD across various systems, each characterized by distinct temporal resolution, such as climate and medical systems like epilepsy or asthma. For medical systems, changes in system states may occur within minutes (i.e., the transformation of the patient from no illness to an illness state), necessitating high-frequency data sampling such as seconds datasets. Conversely, ecological systems such as forests typically exhibit slower dynamics, with state transitions often spanning months or years. The greenness of the vegetation does not change significantly within two consecutive days. Considering the pace of ecological dynamics, working at daily time scale would not add information in terms of resilience analyses. Thus, monthly or bi-monthly sampling frequencies are adequate for calculating early warning signals and capturing resilience dynamics within forest ecosystems, which is aligned with previous studies[10,11]. Overall, we need to select the appropriate sampling frequency based on the specific scientific questions.

(2) Additionally, high-frequency optical satellite data (e.g., daily temporal resolution) frequently suffers from significant noise caused by cloud cover and varying sky-view angles, substantially decreasing data quality[40]. Such noise leads to substantial data loss and changes the signal-to-noise (SNR) ratio of the data, which also diminishes the accuracy and reliability of resilience indicators estimation. The SNR refers to the relative strength of the meaningful ecological signal compared to the noise introduced from observational errors, atmospheric disturbances (e.g., cloud cover), and satellite sensor issues (e.g., varying viewing angles). A lower SNR, typically arising from daily satellite observations affected by frequent cloud contamination, aerosol interference, and variable viewing angles, introduces substantial observational noise. High noise levels obscure the ecological signal, significantly diminishing the accuracy of AR1 estimations[14]. Under such conditions, the AR1 parameter will increasingly reflect random fluctuations rather than true ecosystem dynamics, potentially leading to misleading interpretations of resilience and critical transitions. **Therefore, most current studies do not calculate resilience indicators on a daily scale (as it is unreliable), but rather based on data from 8 days to months[3,10,25].**

(3) The MVC (Maximum Value Composite) method mitigates the negative impacts of low SNR by selecting the highest-quality observations from multiple measurements within a temporal window (e.g., monthly). By aggregating multiple observations, MVC effectively enhances the SNR by filtering out the lower-quality (noisier) data points, thereby improving the reliability of AR1 estimation. Our study's method **aligns with previous studies** applying the CSD frameworks[3,17,19], where monthly MVC-composited data have been effectively utilized. We also presented a global resilience map based on our method (Fig. S2), and **the results were consistent with existing previous studies[10,11,15] (Fig. R6).**

Action Taken. To ensure the robustness of our results, we conducted comprehensive sensitivity analyses regarding temporal resolution. Specifically, we compared the resilience metrics (such as λAR1) calculated using (1) monthly LAI data synthesized by the MVC compositing method and (2) original 8-day LAI data from a single sensor (Terra), both at their native spatial and temporal resolutions. The results show that AR1 values derived from both the 8-day and monthly LAI datasets exhibit highly similar spatial patterns (Fig. S25). We extended these comparisons to other vegetation indices as well, and consistently observed similar patterns in resilience difference across managed forests, regardless of the temporal sampling frequency (Fig. S23~S24). Moreover, analyses restricted to the growing season and additional tests considering data gaps also demonstrate that neither temporal sampling frequency nor missing data has a significant impact on our main conclusions (Fig. S22~S25). Thus, our findings are robust to the choice of temporal resolution and data continuity.



Supplementary Figure S25. Spatial patterns of long-term λAR1-LAI are shown for two regions (the western U.S. and southern Asia, as shown at upper and lower panels, respectively) to illustrate the effect of time-series pre-processing (as described in Figure S22 caption including GS_Single, MVC, and Single) on the resilience estimation.

Moreover, in response to the reviewer's comments, we conducted additional analyses using different temporal resolutions and time spans to evaluate the robustness of our results. Specifically, we analyzed KNDVI data at both 16-day and monthly intervals over four periods: 2001–2010, 2001–2015, 2001–2020, and 2001–2024. MOD13A2 provides 16-days NDVI at 1km from 2001-2024, and MOD13A3 provides monthly NDVI at 1km from 2001-2024. The results indicate that our findings are not sensitive to the length of the time series or the observation frequency (Figure R7 and Fig. S30). The detailed discussion and method are displayed in the supplementary materials (Lines 839~843 and Lines 184~198 in the supplementary materials, i.e., Supplementary Note 5).

?

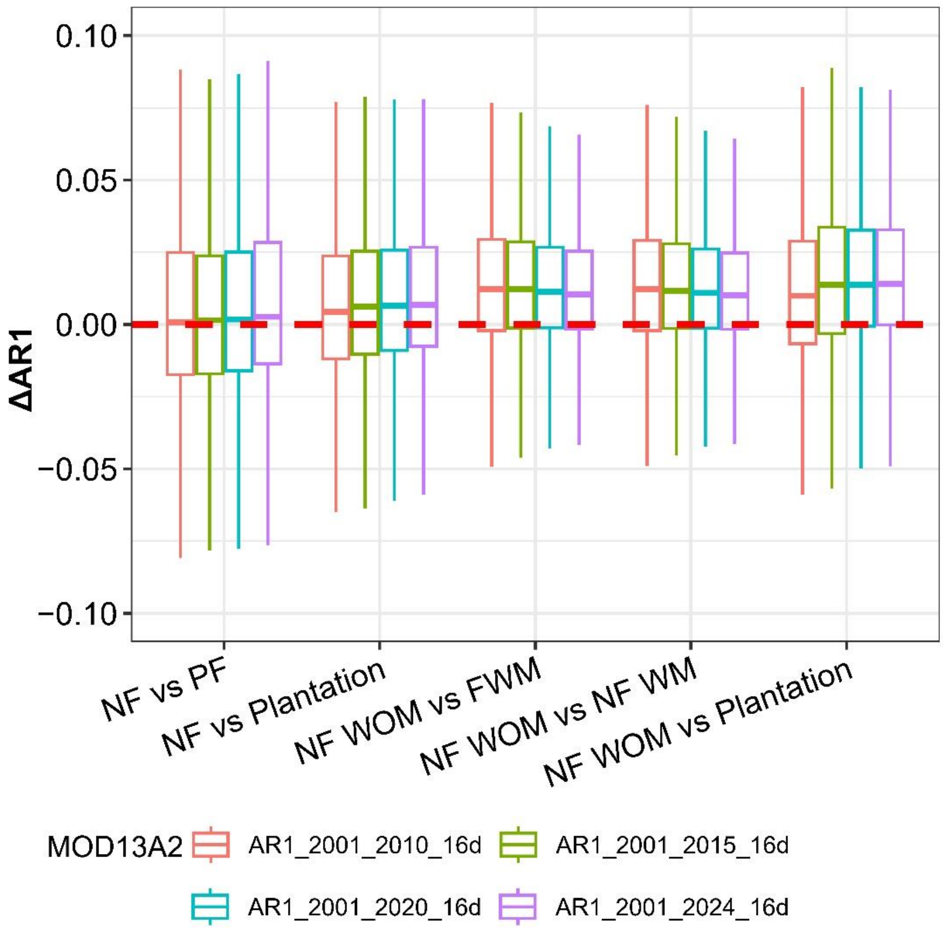
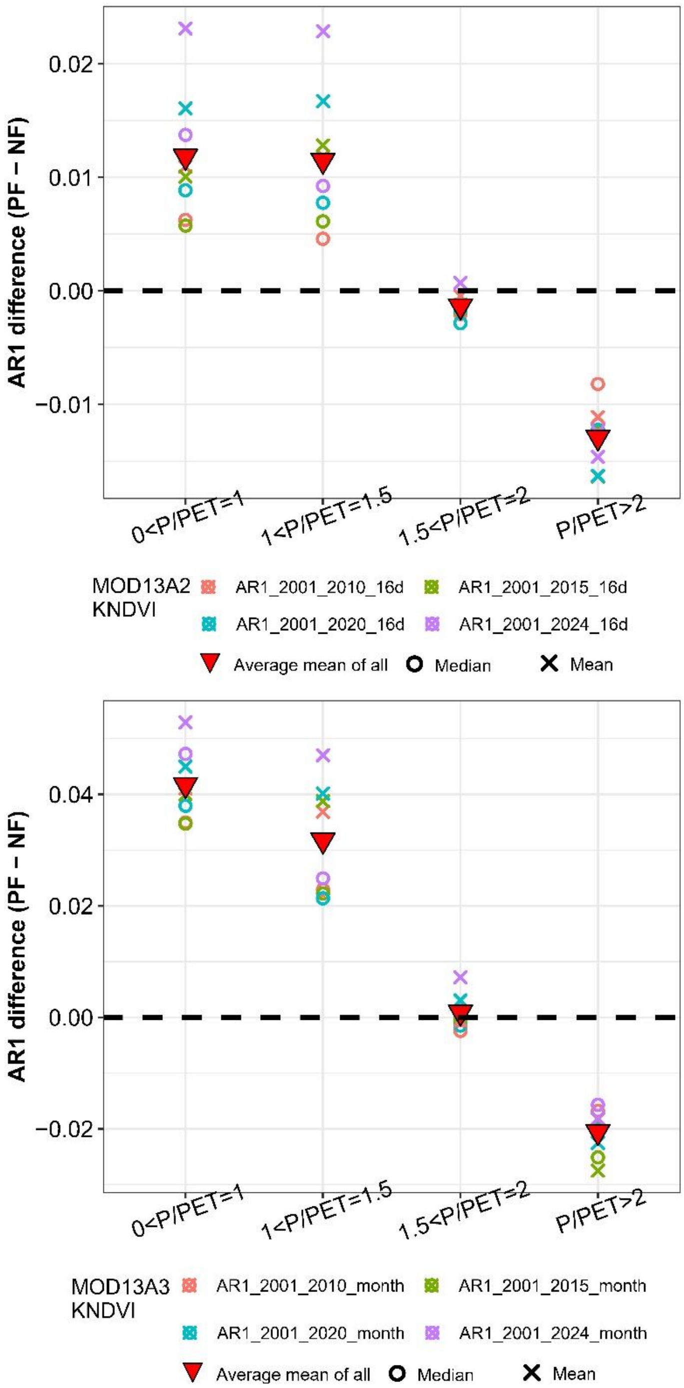


Figure R7. Similar as Figure 2(k), but for ΔAR1 estimated by MOD13A2 datasets (16days is the temporal resolution). Each boxplot represents ΔAR1 derived from time series AR1 indices based on different time spans (2001–2010, 2001–2015, 2001–2020, 2001–2024). Comparisons include NF versus PF, plantation, and various managed/unmanaged conditions. Red dashed line indicates ΔAR1 = 0 baseline. Box colors represent the corresponding AR1 calculation period.



Supplementary Figure S30. Effect of temporal resolution and time spans on resilience differences between PF and NF. Similar to the Figure 3c, but by MOD13A2 (16days, 1km) and MOD13A3 datasets (monthly, 1km). Each point represents ΔAR1-kNDVI $_{(PF, NF)}$ over different time spans (2001–2010, 2001–2015, 2001–2020, 2001–2024). A cross (×) marks the mean of each vegetation index, a circle (○) denotes the median, and a red inverted triangle (▼) indicates the average of the mean and median values.

but then you need to deal with changes in viewed angles and missing data, which I suppose incapacitate the use of these AR techniques. Furthermore, the MVC technique to composite means the actual time stamps of each month is not regularly separated from the former or from the next, so this further means the AR technique is applied in sup-optimal conditions (i.e. non regular timesteps). All these issues are not addressed.

R: The method of MVC can make the time interval of data fixed. Here we will explain the MVC in detail: The Maximum Value Composite (MVC) method for LAI, as utilized in our study, aggregates satellite observations initially provided at 8-day intervals. Within each month, typically three to four observations are available for LAI, and the MVC selects the highest vegetation greenness value, **assigning a standardized timestamp (the first day of each month, e.g., 2010-01-01, 2010-02-01, 2010-03-01)**. Consequently, the temporal intervals between consecutive monthly composites are consistent, each separated precisely by one month.

In rare cases where no valid observation is available for an entire month (due to persistent cloud cover or other observational constraints), this month's data is considered missing. It is important to note that the presence of missing values does not prevent AR1 from being computed. Any lagged temporal matching and such missing data points are excluded from AR1 calculations.

We follow the conventional approach in previous studies[10,15], where missing values are simply ignored and the calculation proceeds with the available data. Other approaches, such as interpolation, may introduce additional uncertainties. For example, interpolation based on vegetation–climate relationships can create dependency issues, while moving averages or low-pass filters may distort high-frequency signals, leading to underestimation of short-term variability and loss of early warning signals. Previous studies have also shown that missing values have limited influence on resilience estimates[15].

Action Taken. To further ensure the robustness of our results, we performed sensitivity analyses (Figs. S22–S25), including comparisons between resilience indicators calculated using only the growing season data and using data from all months. The growing-season-only data contain missing values. We found that the main conclusions remain unchanged, demonstrating that our results are not sensitive to missing data.

I know others have used these metrics on such similar data before (e.g. Forzieri et al.), but that does not convince me either, as I believe the same flaw applies there. I would require a proper rebuttal plus explanation to address these specific issues to accept the premises of this manuscript. Below I further provide more (non exhaustive) specific points to help the authors improve the text.

R: Thank you for your critical comments regarding the theoretical basis of our approach. The use of CSD indicators is grounded in well-established ecological theory, where a loss of resilience is reflected by statistical changes (such as increased autocorrelation and variance) in system state variables. Remotely sensed indices serve as integrative proxies for ecosystem state at broad spatial and temporal scales. As discussed in our previous responses, we have provided a detailed explanation of these points.

To further strengthen the theoretical grounding, we also sought feedback from experts who have applied similar approaches (e.g., Forzieri Giovanni, Taylor Smith) during the revision process, which helped us clarify the conceptual link between the CSD framework and forest ecosystem dynamics in this study.

Specific points

L73: Management should not be plural. Perhaps change to "types of management"

Action taken. We have checked the usage of singular form of “management” throughout the manuscript to ensure its correctness. We have corrected this text in the revised version. The modified version of the sentence now reads (Lines 57~61):

“Recent studies have used process-based models to predict forest resilience with moderate accuracy in certain regions[22], but their global predictions still exhibit considerable uncertainty[23], particularly under complex environmental conditions and in areas exposed to forest management, such as timber harvesting, reforestation, afforestation, and thinning[24].”

L82: CSD should be spelt out here

Action taken. We have corrected the text in the revised version of the manuscript according to the reviewer’s suggestion. The modified version of the sentence now reads (Lines 74~75):

“Within this context, the Critical Slowing Down (CSD)[8,30–32] theory has been formulated to quantitatively measure ecological resilience.”

L91: These new metrics should be explained, even if briefly, not just cited. The reader should get here the diffrence between AR1 and lambda_AR1. What does the lambda stand for? What is Lambda variance?

R: Smith's study discovered a nonlinear relationship between AR1 and the ecological resilience (estimated by the recovery process after any disturbances). Subsequently, they proposed λAR1 based on the nonlinear relationship and the O-U process (similar for Variance as well).

Action taken. According to the reviewer’s suggestion, the modified version of the sentence (explaining the relation between AR and λAR1) now reads (Lines 78~81):

“Such loss of resilience can be reflected by CSD-based early-warning signals, such as increases in 1-lag autoregression (AR1) and variance (Var), as well as their revised formulations (λAR1 and λVar) that account for the underlying non-linear relationship with ecological resilience[16,34–36].”

Further details on the differences and relationship among these indicators and their equations can be found in *Method section* (Lines 639~650 and Lines 328~355 in *Supplementary Note 9*).

L94: what is "variance" by itself? you need to be more specific... the temporal variance of the signal... and over how long of a period?

Action taken. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We have moved the related part into the Methods section. Here, we mention the resilience indicators we calculated, where the long-term average values are derived from 2001–2015, and the temporal changes are assessed using a 4-year moving window. The modified version of the sentence now reads (Lines 656~663):

"Forest resilience is quantified as a long-term ecological characteristic and derived from the CSD indicators computed over the whole time series (2001–2015) at each pixel (1-km) (e.g., 2001 – 2015 λAR1-LAI). This estimate describes the long-term average resilience of forest ecosystems. We complement this analysis by exploring the temporal changes of forest resilience, deriving CSD indicators over 4-year moving windows, as proposed in recent studies[35]. This results in a time series of CSD values over the period 2003 – 2013 at each pixel. Resilience dynamics are expressed as linear trend of CSD indicators (e.g., λAR1-LAI trend)."

L97: In this paragraph you speak of statistical metrics, but you should have also mentioned the signal on which the metrics need to be applied, otherwise it is pointless. They cannot be applied to a signal that does not speak about resilience. So first mention how RS signals can effectiveely (or not) be sensitive to properties that speak to resilience, and only then introduce the statistical metrics.

Action taken. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We have moved the related part into the Methods section. We first introduced the de-seasonality and de-trendiness of remote sensing indices, and then presented different CSD indicators based on these signals (Line 627~650).

In addition, for readers' understanding, we briefly mentioned the relevant steps at the beginning of the Results section (Line 99~114).

L104: Not clear why ML comes into play? Why would it be needed? needs to be introduced. Also the "impacts of biophysical factors". What are these "biophysical factors"? L110: "biophysical factors" should have been mentioned and described before

R: **Action taken**. In the revised text, we have clarified the rationale for introducing machine learning. Specifically, we aim to quantify the relative contributions of various biophysical factors to resilience differences, recognizing their potential nonlinear effects. The modified version of the sentence now reads (Lines 94~96):

"Furthermore, we employ machine-learning techniques to elucidate the non-linear interactions between biophysical factors and resilience difference across forest management types."

We have also provided a clear definition of the term "biophysical factors" for the first time, which include climate variables (e.g., P/PET, temperature), soil characteristics (e.g., SOC), and vegetation variables (e.g., root depth). The modified version of the sentence now reads (Lines 85~87):

"In this study, we investigate how different forest management practices interact with biophysical factors (e.g., climate and vegetation characteristics) and explore their joint effect on forest resilience globally during 2001–2015."

The reasons about the selection of these variables are displayed in the Methods section (Lines 759~760):

"These factors, including key climatic, vegetation and soil characteristics, are selected for their documented influence on forest resilience[17,36]."

L117: why tree cover from 2000 and management for 2015? the former is available for 2000

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. We have corrected the use of prepositions, now reading as (Lines 99~102):

"We derive the distribution of forest management types for 2000 (GFMD 2000) as the reference map (Fig. 1a). We generate this by overlaying tree cover data from the Global Hansen Forest Change Map for 2000 (2000 GHFC) with the Global Forest Management Datasets for 2015 (2015 GFMD)."

In addition, we have also corrected the related description in the method section (Line 554~563).

L122: kNDVI is with lowercase k

R: **Action taken**. We have corrected the full texts in the revised version of the manuscript according to the reviewer's suggestion.

L127: again, the acronyms are meaningless and useless if we do not know what they measure

R: **Action taken**. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The modified version of the sentence now reads (Lines 110~112):

"In this study, forest resilience is measured by three CSD indicators globally: AR1 (lag-1 autoregression), λAR1 (lambda lag-1 autoregression, Fig. 1e), and λVar (lambda variance)."

L499: Here is a line that puzzles me. If we need stationary behaviour, and "abrupt events and human disturbances lead to non-stationary conditions"

R: Here, our intention was to emphasize that CSD indicators should be computed from data that are stationary, meaning that trends and seasonality should be removed beforehand.

Some studies suggest that the increasing frequency of extreme events and human disturbances may induce non-stationary conditions in the future53, potentially undermining the reliability of CSD-based early warning indicators. After removing seasonality and trends, the time series is expected to remain stationary. However, under future scenarios, the increasing frequency of extreme events may disrupt this stationarity assumption, thereby undermining the reliability of CSD-based early warning indicators. Thus, applying the CSD framework to future projections requires caution (e.g., the application of CSD in CMIP6). Previous study have highlight that there is considerable uncertainty about how resilience will change in the future by CMIP6 outputs35.

But applying CSD indicators to current remote sensing vegetation data remains robust and widely accepted[12].

Action taken. In response to your suggestion, we have deleted the related sentence.

L502: This scale dependence is crucial to take into account in my opinion as in this case, natural forest will likely have different (smaller) spatial extents of perturbations than managed forest (where, by definition, there would be more clear cuts).

R: We thank the reviewer for this important observation. We agree that differences in the spatial extent of perturbations between natural and managed forests are critical when interpreting resilience indicators derived from remote sensing data.

To address this concern, we conducted sensitivity analyses on both temporal and spatial resolutions. For temporal resolution, we compared CSD-based indicators (e.g., AR1) derived from 8-day and monthly LAI time series across different forest management types. The results showed consistent resilience patterns, indicating that temporal resolution had little impact on our conclusions (Fig. S22~S25, Fig. S30). For spatial resolution, we used both 500 m and 1 km LAI data (with 1 km data, which were generated by aggregating the 500 m data) to estimate resilience indicators (Fig. S29). We found that the spatial resolution did not significantly alter the observed resilience differences between planted and natural forests.

In addition, we tested the paired-grid method, and thus the assessment of the resilience difference between forest management types, with respect to varying disturbance regimes and ecological contexts. For example, we constrained the differences in forest cover change between PF and NF to 3–10%, and the differences in AGB to 5–10 Mg/ha. We found that the resilience differences still exhibited similar patterns (Fig. S21).

Action taken. Furthermore, we discuss this limitation in the Discussion section, now reading as:

"Finally, we examine dataset-related factors, such as classification accuracy (Supplementary Fig. S27-S28), spatial resolution (Supplementary Fig. S29), and temporal spans and frequency (Supplementary Fig. S30). Overall, these experiments show that our results are minimally affected by such methodological definitions and assumptions (details in *Supplementary Notes 3-5*)." (Lines 460~465)

"The resilience modeling presented in this work is grounded on the CSD theoretical framework. CSD-based indicators (AR1, λAR1, and λVar) have been widely used in large-scale applications to explore early-warning signals of critical transitions in forest ecosystems[16,17]. However, these metrics may be scale-dependent, as ecological processes vary across spatial and temporal scales[35]. We have evaluated this issue by replicating our resilience assessment for different spatial and temporal resolution of the input data and found a substantial agreement amongst the different experiments conducted (Supplementary Fig. S29-S30)." (Lines 482~489)

L508: GPP is not a biomass index. And the way it is calculated (from MODIS) is very likely affected by the same saturation effect of VIs.

R: We agree that both LAI and GPP may be subject to saturation effects, particularly in high-density forests, and thus may introduce uncertainty in resilience estimates. To mitigate these limitations, we included the kNDVI index in our analysis, which has been shown to be less sensitive to saturation in dense canopies. The patterns observed using kNDVI are consistent with those from other optical indices, supporting the robustness of our results (see Supplementary Materials for details). Unfortunately, true biomass indices such as VOD have very coarse spatial resolution (~0.25°), making them unsuitable for fine-scale analysis of forest management effects in this study.

Action taken. We have corrected the text in the revised version of the manuscript according to the reviewer's suggestion. The modified version of the sentence now reads (Lines 491~500):

"We recognize that spectral greenness metrics employed here, such as LAI, are prone to saturation in dense forests, an effect that can potentially bias resilience estimation[35]. Alternatively, vegetation indices based on microwave sensors, such as Vegetation Optical Depth (VOD), can be valuable alternatives; however, their coarse spatial resolution (e.g., 0.25°) limits their ability to detect fine-scale management-related patterns. To test the robustness of our findings with respect to possible saturation effects, we employ kNDVI, a recently developed spectral vegetation index that has been documented to be more closely related to primary productivity and resistant to saturation compared to similar products based on surface reflectance of vegetation[51]."

L614: MVC only makes sense for NDVI, not for (downstream) other products (e.g. LAI, fAPAR, GPP, etc.), so be precise on whether you use it only for NDVI or not, and what are you using for the other signals. Also, the reference of MVC should be the original, Holben et at.

R: We thank the reviewer for this important comment. While MVC was originally developed for NDVI (Holben et al., 1986), its use has been extended to other vegetation-related products such as LAI, FPAR, and GPP in numerous studies[54–56]. Products such as LAI also have problems such as cloud interference and inversion errors (similar to NDVI products). For example, previous studies applied MVC to LAI products to generate high-quality time series for vegetation monitoring.

In our study, we used the MVC method for all vegetation indices to ensure the robustness of the derived resilience indicators (e.g., λAR1), especially under variable quality conditions. The use of monthly maximum compositing allows us to retain the clearest signal of vegetation activity while mitigating the effects of noise and missing values. Meanwhile, we also referred to other studies that applied the CSD theory to monthly vegetation indices.[12,19]. **Action taken**. The original literature citation has been replaced by Holben et al. (1986).

L668: doesn't this window width also depend on the frequency of observations?

R: We thank the reviewer for raising this important point. The window width was essentially chosen based on recommendations from previous studies. In addition, we also tested the effect of different sampling frequencies under a fixed 4-year window. For monthly MVC data, this corresponds to a window width of 48 observations (12 × 4 years). For datasets with a 16-day interval, the same four-year window typically includes 92 observations (23 × 4 years). These two cases correspond to the MVC-based kNDVI data and the Terra-only kNDVI data used in our analysis, respectively (Fig. S22~S25). The results show that the sampling frequency has a relatively small influence on the resilience estimation.

In addition, similar sensitivity tests about the frequency can also be found in the sensitivity test section (Lines 839~843 and Fig. S30).

L672: Why MVC?

R: **Action taken**. In the revised manuscript, we have clarified the rationale for using the Monthly Maximum Value Composite (MVC) method. The MVC approach is widely applied in remote sensing studies to minimize atmospheric contamination and view angle effects, thereby improving the consistency and reliability of vegetation time series used in resilience analysis (Lines 616~624).

L674: you should be more specific, as here what you are doing (with MVC), is to downsample the frequency of actual observations. if the monthly values are effectively a single value in the time series (as with MVC), the time between consecutive observations is not a month, but van vary. So what does that do for CSD?

R: To clarify, we applied the Monthly Maximum Value Composite (MVC) method by selecting the maximum value of vegetation indices within each calendar month to represent the clearest vegetation condition for that month. After compositing, each monthly MVC value is assigned a consistent time stamp (i.e., the 1st day of each month), thereby ensuring that the resulting time series maintains a fixed monthly interval. Although the actual observation dates of the selected maximum values may vary within the month, the time series used for Critical Slowing Down (CSD) analysis is temporally regular, with one observation per month (i.e., LAI-value on 2020.01.01, LAI-value on 2020.02.01, LAI-value on 2020.03.01, …).

Action taken. We also performed sensitivity tests to ensure that this compositing approach does not bias our key results. The impact of potential temporal irregularity introduced by MVC selection was found to be minimal, with resilience patterns remaining robust across different temporal resolutions and compositing methods (Fig. S22~S25).

Considering your suggestion, we have modified the sentence to be more specific (Lines 616~624): "Furthermore, we only retain high-quality pixels, according to the quality flag of the specific products, and synthesize monthly data using the maximum value composite (MVC) method[63]. Specifically, for each calendar month, the maximum value of the vegetation index (e.g., LAI) is selected from all available observations within that month (e.g., from both Terra and Aqua). Each composite value is then assigned to the first day of the corresponding month, resulting in a temporally regular monthly time series, regardless of the actual observation date of the maximum value."

L690: How good is the Besnard et al data to differentiate forest age data over a 0.25 by 0.25dd area? it is a requirement that this is very high if this is used in a moving window approach, otherwise this does not provide robust info

R: We thank the reviewer for this important question. While the moving window in our study is performed at a 0.25° × 0.25° spatial resolution, the forest age data we used from Besnard et al. (2021) are originally provided at a much finer 1 km resolution (available at: https://www.bgc-jena.mpg.de/geodb/projects/FileDetails.php). Here is the forest age map by 1km (Fig. R8), which shows that there are very clear differences (i.e., high spatial heterogeneity).

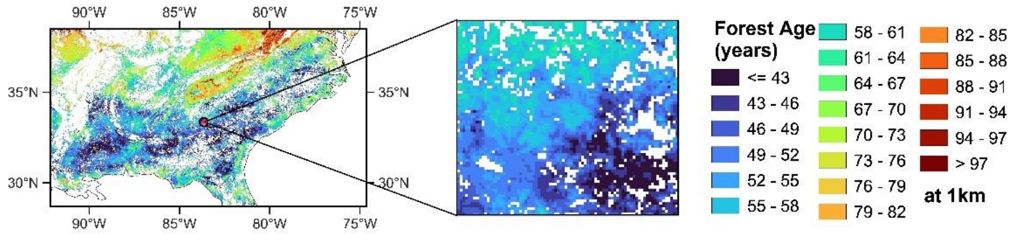


Figure R8. The spatial distribution of forest age data.

Action taken. We have corrected the text in the revised version of the manuscript by adding the spatial resolution (Lines 687~689), now reading as: "The global forest age data at 1 km is obtained from Besnard et al.[67], while elevation data at 30 m is acquired from the Shuttle Radar Topography Mission (SRTM) v4 dataset[68]."

The following are 3 other specific points regarding the response to reviewers.

1

Figure 1. Action taken. (2) Background grid lines have also been added to subplots b, d, and e to aid comparison (add the Lat/Lon in the figures).

I do not see any background grid lines. Maybe they were lost due to the low-resolution of the figure?

R: In our R1 response, we mistakenly described the modification. What we intended to convey is that, compared to the previous version, we have added latitude and longitude axis labels to the subplots, so that the specific locations can be identified. We chose not to add background grid lines (latitude/longitude lines) directly onto the figure to avoid obscuring key spatial details. If you have specific suggestions on how to improve the visualization, we are happy to make further adjustments. Additionally, we will increase the resolution of the figure and provide a high-quality PDF version to the editors for publication.

2

L168-169: "over 40% of paired grids show higher resilience in PF than in NF" this is quite high! If this is the case, I wonder if the statements about management decreasing resilience is founded or whether you should be more conservative. Considering the relevant uncertainties, is the 60%-40% ratio really different than a 50%-50% ratio based on random noise?

R: Thank you for this important question. Our intention is not only to highlight that certain management activities can diminish resilience, but also to emphasize the complexity and variability of these effects. While forest management activities associated with high human pressure—such as intensive plantations—can significantly reduce forest resilience, there are also management practices characterized by lower anthropogenic pressure that may enhance resilience under specific conditions. For instance, as illustrated in Figure 2, the most intensive form of human planting (i.e., plantations) shows the greatest negative impact on resilience, whereas less intensive afforestation practices (i.e., planted forests) have a relatively weaker effect on reducing resilience. Action taken. Nonetheless, we recognize that the reader may be misled here, so we have rephrased the paragraph topic, now reading as (Lines 184~186). "Although forest management activities with strong human pressures have weakened forest resilience, managements with lower anthropogenic pressures can enhance resilience under specific conditions (Fig. 2)."

Not sure if this is answered. My original comment referred to the fact that the percentage of pixels where one or the other conclusion could be drawn is potentially indistinguishable from random noise. The authors state that “40% of paired grids show higher resilience in PF than in NF” (L186-188). This means that the main outcome is that 60% of paired grids show lower resilience in PF than in NF. Given the many confounding factors affecting natural systems, how confident are you that these results are different than a null hypothesis where 50% of paired grids show higher resilience in PF than in NF, and 50% show lower resilience in PF than NF? While I agree with the explanation provided by the authors about the management effects on forests, this misses to answer the raised concern about the percentages and random noise, which is a statistical issue.

R: Thank you for raising this important statistical point. We agree that it is crucial to determine whether the observed 60%–40% ratio truly deviates from a 50%–50% distribution that would be expected under a null hypothesis of no difference between PF and NF resilience.

To formally assess this, we conducted binomial tests on the proportion of paired grids where PF showed lower resilience than NF. The results were statistically significant ($p < 0.05$) for several vegetation indices, including kNDVI and GPP_PML, suggesting that the observed pattern is unlikely to be due to random variation alone (see Fig. 2). We also acknowledge that not all indices showed statistically significant deviations—for example, LAI, fAPAR, and GPP_MODIS did not yield significant differences from a 50% distribution. However, when we evaluated λVar as an alternative indicator of resilience, we found that all four indices (except kNDVI) exhibited a statistically significant tendency for PF to have lower resilience than NF (see Fig. S3).

Taken together, these multiple lines of evidence support a general but index-dependent trend: planted forests tend to exhibit lower resilience than natural forests. In the following text (Lines 196~218), we further indicate that P/PET can alleviate the difference in resilience between artificial forests and natural forests, and even reverse the pattern.

3

Supplementary Figures. Given the updated figure specifications (e.g. color scales), try to also update the supplementary figures to be consistent with the figures in the main text. E.g.: the maps in Figure 2 and Figure S4 have now a different color bar.

R: **Action taken**. We have corrected the color (to red-blue color) in the revised figures of the manuscript. (see Fig. S4 and S7)

Response to Reviewer #2

Reviewer #2 (Remarks to the Author):

I co-reviewed this manuscript with one of the reviewers who provided the listed reports. This is part of the Nature Communications initiative to facilitate training in peer review and to provide appropriate recognition for Early Career Researchers who co-review manuscripts.

R: We greatly appreciate the time and effort you have dedicated to reviewing our manuscript and providing important feedback. We have modified and responded point-by-point accordingly.

Response to Reviewer #3

Reviewer #3 (Remarks to the Author):

The authors have adequately addressed my primary concern regarding the effect of age on the main conclusions of their study. Additionally, they have satisfactorily resolved other issues raised in my initial review. After carefully rereading the revised manuscript, I find no further comments or concerns.

R: We sincerely thank the reviewer for the positive feedback and for acknowledging our efforts to address the previous concerns. We are grateful for your valuable time and constructive input, which helped improve the quality of our work.

Response to Reviewer #4

Reviewer #4 (Remarks to the Author):

1.Comments for authors

The focus of the study on forest management types, biophysical factors and management practices, and forest ecosystem resilience" is timely and interesting, as the results will help to inform decision-making on forest management best practices for various forest management types to build resilient forest ecosystems in a changing climate.

I took the time to review the author response to all the reviewers' comments, supplementary materials and main document, and I think the authors have greatly improved on the manuscript in terms of technical and scientific methodology and clarity in the presentation of the results. In particular, the classification of the forest management types by the authors is much more improved and clearer, enabling better comparison between forest management types. The conceptual figure 2(I) that illustrates the decreasing vegetation resilience across the different forest management types due to human impact is much appreciated. The objectives of the study are clearer and are well answered.

The role of forest management practices (e.g., logging rotation cycle, restoration efforts - for managed natural and planted forests) and some biophysical factors in shaping the resilience of forest types/ecosystems has been addressed in greater detail in the text and figures, which has significantly strengthen the ecological context and applicability of the findings.

Furthermore, the authors have provided clearer explanations of the models and their validation procedures, and essentially specified how the models were assessed, including performance metrics, validation techniques, and potential sources of bias and uncertainty in the spatial data. Therefore, I agree with the methods used and results reached by the authors.

R: We appreciate you taking the time to review our responses to all the reviewers in the previous round and give us encouraging comments. We are grateful for your recognition of the improvements made to the manuscript, particularly regarding the classification of forest management types, bio-physical factors related to vegetation resilience under different management. Furthermore, we have point-by-point replied to the other comments you raised here.

2. However, I still have some major comments/suggestions

Major comments

I think that the results are mainly confirmatory of exiting factual knowledge and merely qualitative in the way that they are presented in the text. I think that the results could be better quantified, precisely when comparing the resilience between forest management types. In particular, instead of using "higher" to qualify unmanaged natural forest's superior resilience over managed forests and planted forest, for example, the authors could quantify the resilience difference in terms of how much divergence or variation that the study found (resilience difference) that may differ from previous global studies. For example, how many times or by how many folds is one forest management type more resilient than the other? This is an important quantification information that provides more context for comparison than merely using the words "higher" or "greater" and "lower" resilience to express the resilience difference between forest management types or forest ecosystems. For example, if an unmanaged natural forest is subject to intensive management practices, how many folds will be the decrease in the resilience factor? This is because, it is factual knowledge that unmanaged NF are older and intact, and thus, are more resilient than managed NF and PF in all climatic and biogeographical environments, and what this means in the context of climate change. Thus, providing the magnitude of the resilience difference in terms of numbers (number of times or folds resilience increased or decreased between forest management types due to forest management practices) should be one of the research gaps to be filled in the existing literature. As a demonstration, in forest carbon management studies that compare natural and regenerating forests, for example, results are presented in a way that is more captivating such as "old growth (unmanaged) natural forest sequester 10 to 20 times more carbon than young secondary or regenerating forest (managed forest)"; is such quantitative comparison possible with the current study between forest management types or forest ecosystems?

R: Thank you for this valuable suggestion. We agree that expressing resilience differences quantitatively rather than using purely qualitative descriptors such as "higher" or "lower" can greatly enhance the interpretability of our findings.

Action taken. In response, we have also correlated the main texts according to your suggestions (i.e., more quantitative description). We also calculated the ratio of resilience between managed forests and natural forests to understand the relative difference between the two (Fig. S3). Here are some modified texts, now reading as:

"Consistently, resilience in NF WOM is approximately 10% higher than in NF WM and about 20% higher than in Plantations (Supplementary Fig. S3)." (Lines 160~161)

"When P/PET is less than 1.5, the resilience of PF (e.g., $\lambda AR1\text{-}LAI_{(PF)} = -1.45$) is weaker than that of NF (e.g., $\lambda AR1\text{-}LAI_{(NF)} = -1.60$). Conversely, when P/PET exceeds 1.5, the resilience of PF (e.g., $\lambda AR1\text{-}LAI_{(PF)} = -2.31$) surpasses that of NF (e.g., $\lambda AR1\text{-}LAI_{(NF)} = -2.15$) (Fig. 3f, g and Supplementary Fig. S5)." (Lines 203~206)

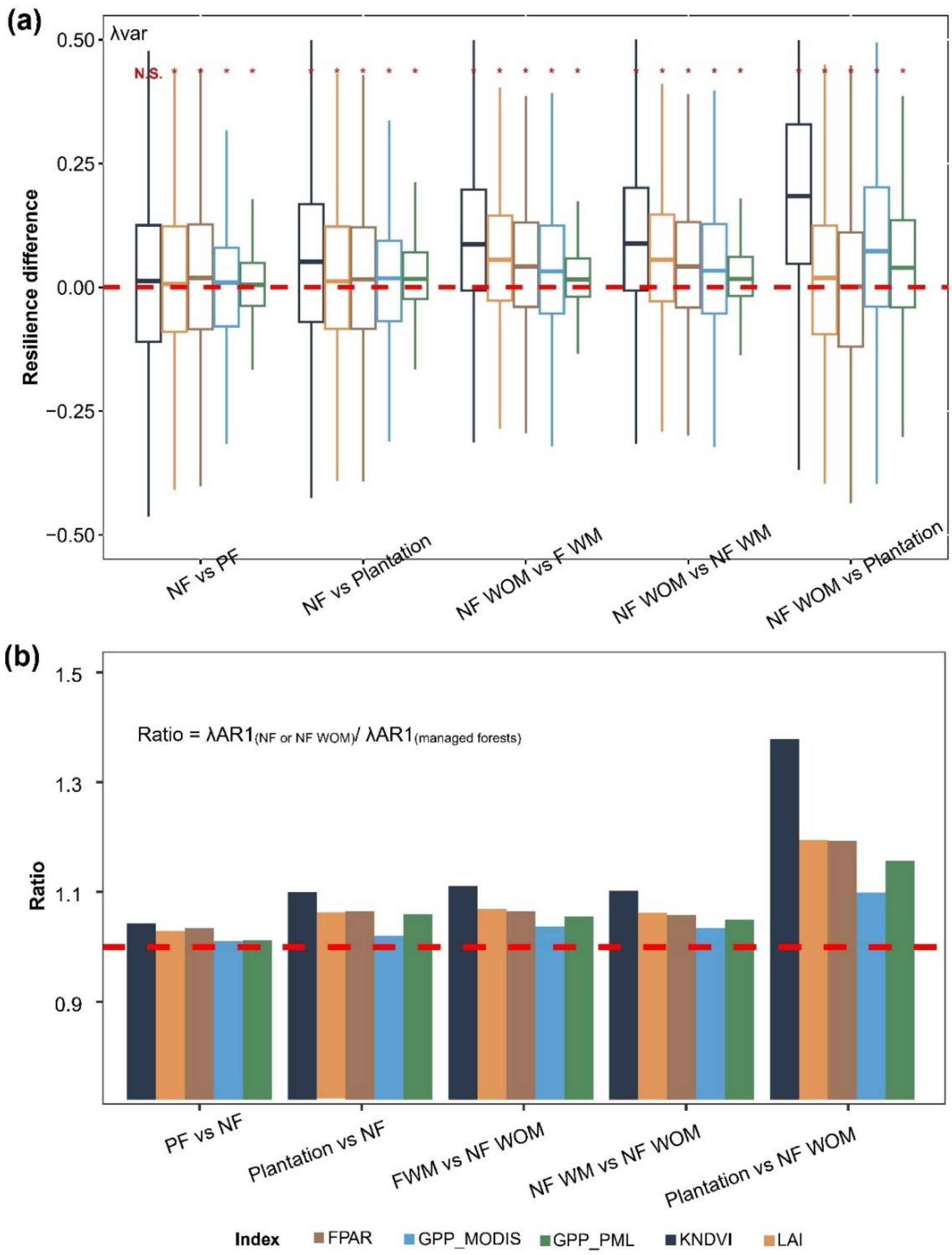
"In agreement with this pattern, results indicate that NF WOM exhibit resilience approximately 13% higher than that of NF WM, and about 20% higher than that of Plantations (Fig. 2k and Supplementary Fig. S3), reflecting again that stronger human intervention can largely reduce forest resilience." (Lines 398~401)

Here we also added the new discussion and supplementary materials to relate the resilience difference and carbon storage, now reading as:

"This resilience difference is key to evaluating the capacity of forest management practices in ensuring the stability of ecosystem functions and services, such as climate mitigation and water regulation[5]. Indeed, their provision is closely dependent on the capacity of forests to withstand and recover from external perturbations (more details in *Supplementary Note 1*)." (Lines 401~405)

"**Supplementary Note 1: The impacts of forest management on resilience and ecosystem services.**

Our analysis indicates that natural forests without management (NF WOM) have 13% higher resilience than natural forests with management (NF WM) (Supplementary Fig. S3). This suggests that forest management may reduce the resilience of natural forests by about 13%. It should be noted that this 13% difference represents an average effect globally. Under specific conditions, such as in dry climates or areas with deep root systems, the reduction in resilience caused by management can become much larger, reaching two to three times the average loss. This pattern aligns with findings from previous studies on forest ecosystem services. For example, unmanaged forests are reported to store two to four times more carbon than managed forests[1–4]. In some regions, this disparity can exceed **tenfold**, i.e. NF WOM may store more than ten times as much carbon as NF WM. Moreover, management practices negatively impact other ecosystem services, including water regulation[5–7] and soil fertility[8,9]. These impacts become even more pronounced under intensive management, such as in Plantations. Compared to NF WOM, the deficits of Plantations in soil fertility, carbon storage, and water regulation are even larger than those observed between NF WM and NF WOM[10]. Their resilience follows similar patterns: Plantations exhibit about 20% lower resilience than NF WOM, which is a greater loss than the loss between NF WM and NF WOM (Supplementary Fig. S3). These observations confirm that ecosystem resilience is tightly linked to broader ecological services. Lower resilience reduces the ability of forests to buffer disturbances, ultimately weakening forest services and ecological stability[11]." (Lines 2~22 in supplementary materials)



Supplementary Figure S3. (a) Boxplots of resilience differences (ΔλVar) between different forest management types, including NF vs PF, NF vs Plantations, NF WOM vs F WM, NF WOM vs NF WM, and NF WOM vs Plantations. Results are shown for five vegetation indices (FPAR, GPP_MODIS, GPP_PML, kNDVI, and LAI). Asterisks indicate significant differences ($p < 0.05$), while N.S. indicates non-significance. The red dashed line represents zero difference. Statistical significance tests (Wilcoxon test) are performed to assess whether these resilience differences are statistically distinguishable from zero. (b) Ratios of λAR1 between different forest management types. The "Ratio" is calculated as the λAR1 of NF (or NF WOM) divided by the λAR1 of managed forests. The red line denotes a ratio of 1, and the bars represent the mean values. Ratio values greater than one show that λAR1 of natural forests is higher than that of managed forests, indicating lower resilience of managed forests.

Minor comments

3. The title is still unclear in terms of the structural presentation of the independent and dependent variables and the interrelationship. The biophysical factors and forest management practices determine the resilience of forest management types. The focus of the study is on the resilience differences of forest management types as influenced by biophysical factors and forest management practices? I suggest that the authors should provide a more thoughtful and informative title.

R: We sincerely appreciate the reviewer's thoughtful comment regarding the clarity and structure of the title. **Action taken**. Based on your suggestion, we modified the title of this study, now reading as:

"Biophysical factors and management practices are key to shaping forest resilience"

If you have a preferred option or would like to suggest an alternative title that better captures the study's focus, we would be very happy to consider it.

4. In the abstract, natural forest should be used instead of native forest, as this is the terminology used throughout the manuscript; the efficient "supply" or "provision" (choose one term) of ecosystem services should be used instead of "functioning"; external disturbances (i.e. ecosystem resilience) is unclear, since external disturbance does not mean the same thing as ecosystem resilience; the former influences the latter.

R: **Action taken**. We have corrected the full texts in the revised version of the manuscript according to the reviewer's suggestion. We have reviewed the entire manuscript and replaced the relevant terminology. We re-write the sentences for better introduction the topic of this study (Lines 5~10).

"Forest management influences key ecosystem services essential for societal well-being. However, how these human activities affect forest resilience (i.e., the system's ability to withstand and recover from perturbations) under different environmental conditions remains poorly understood. This study investigates how different forest management practices interact with biophysical factors and explore their joint effect on forest resilience globally during 2001-2015."

The authors should categorically state in the abstract that unmanaged natural forest demonstrated higher (quantified in terms of magnitude of resilience difference) resilience than managed natural and planted forests (or all the other forest types). The authors clearly stated this result in the text as follows "… the results by the nearest neighbor method showed that unmanaged NF demonstrate higher resilience than managed", but not in the abstract. This is the key result and should be stated in the abstract first, before stating how these may vary under different biophysical conditions such as wet climate (water availability/deficiency modulating the resilience difference effect of forest management types). I think that unmanaged natural forest is more resilient under all climate conditions, though I agree that adapted best management practices can improve resilience under favourable biophysical conditions.

R: **Action taken**. We have corrected the full texts in the revised version of the manuscript according to the reviewer's suggestion. Here are the reversed texts, as reading (Lines 12~16):

"Results show that natural forests (NF) without management exhibit the highest resilience among all forest management types, followed by natural forests with management and planted forests (PF). Most forest management practices weaken resilience, with greater losses observed under more intensive management."

According to suggestion, we have revised the full abstract accordingly to ensure consistency in the presentation of logics that your mentioned.

In the abstract, "These differences are influenced by biophysical contexts such as local climate, vegetation, and soil". However, I think the authors should add one sentence on the limitation of the study, especially in terms of the limited variables included in the model in the abstract.

R: **Action taken**. We have corrected the full texts in the revised version of the manuscript according to the reviewer's suggestion. Here are the reversed texts, as reading (Lines 23~24):

"This study also highlights that more variables (e.g., biodiversity) should be considered for further exploring forest resilience."

We also mentioned this limitation in the Discussion section (Line 509~513).

"Future research should focus on the integration of additional factors such as landforms and species diversity (including richness, evenness, and composition) into modeling to further improve our understanding of how management practices affect forest resilience under varying environmental conditions."

The full stop (.) mark comes after the quotation (") mark.

R: **Action taken**. We have corrected the full texts in the revised version of the manuscript according to the reviewer's suggestion.

There are too many abbreviations and acronyms in the text, please reduce them to a minimum to improve on the readability of the manuscript. For example, all the forest management types could be written in full in the text.

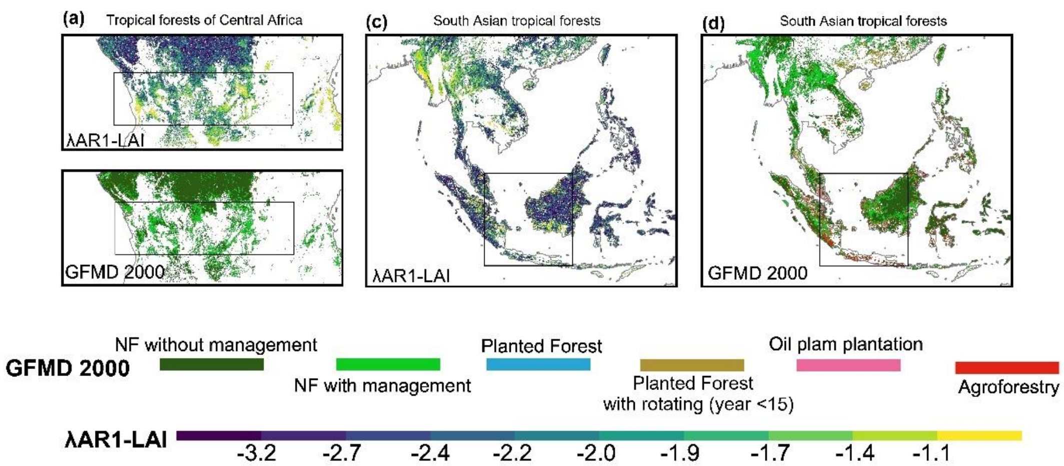
R: **Action taken**. We agree that excessive use of abbreviations may hinder readability. In response, we have ensured that all forest management types are written in full at their first appearance in each major section of the main text and supplementary materials. However, due to the word count limits of Nature Communications, we retain abbreviations after first definition to keep the manuscript concise and within the required length. We have also included a summary table listing the definitions and descriptions of all abbreviations used (Table. S1).

There are still some issues with grammatical accuracy that need greater attention and revision before publication. The authors can seek the help of a native English speaker colleague in the same field of research.

R: We appreciate the reviewer's suggestion regarding grammatical accuracy. We sought assistance from a native English speaker with expertise in the field (Giovanni Forzieri, University of Florence and European Commission) to ensure that the language meets publication standards. With his help, we have revised the entire manuscript to improve the clarity and presentation of the results.

Figure S1. A clearer font (without text effects) should be used to label the forest management types, a single font should be used for all labels in all the figures.

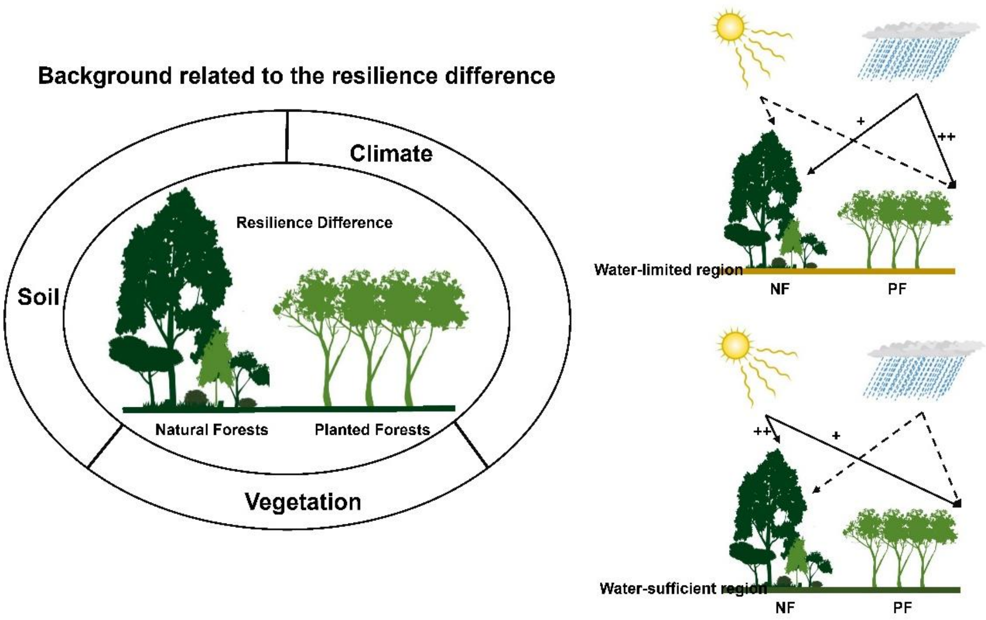
R: **Action taken**. We have corrected Figure in the revised version of the manuscript according to the reviewer's suggestion. Here is the modified figure:



Supplementary Figure S1. The impacts of management on resilience in tropical forests. (a, b) Human management affects the resilience of tropical forests in Central Africa, resulting in lower resilience in natural forests with management (NF WM) compared to natural forests without management (NF WOM, i.e., pristine forests). (c, d) The resilience of oil palms in Indonesia is lower than that of NF WOM.

Figure S12. The shadows in the labels or lettering should be remove.

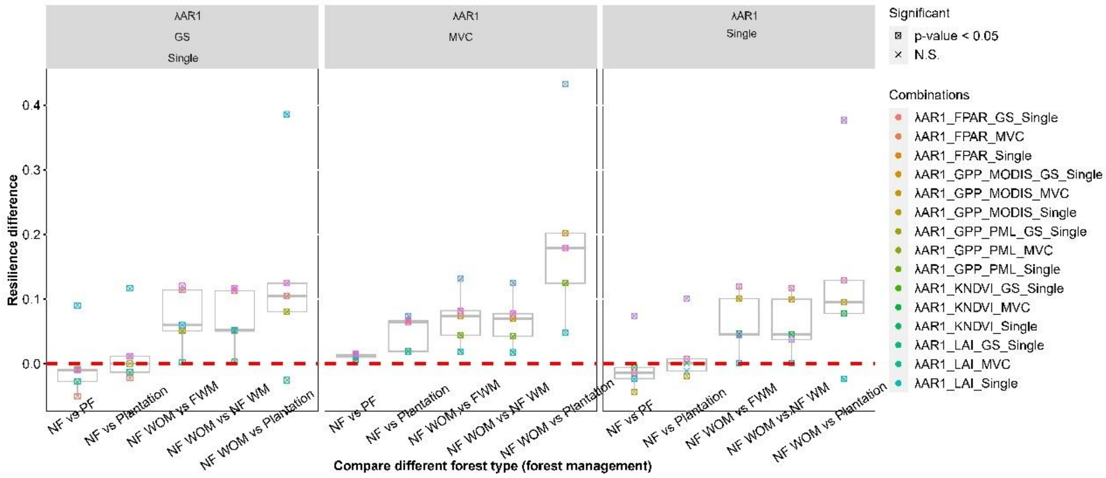
R: **Action taken**. We have corrected Figure in the revised version of the manuscript according to the reviewer's suggestion. Here is the modified figure:



Supplementary Figure S12. Conceptual diagram illustrating how resilience differences between PF and NF are shaped by local biophysical contexts, e.g., climate regions. In water-limited regions (dry areas), vegetation is highly sensitive to water (e.g., soil moisture and precipitation). PF show higher sensitivity to water stress than NF. This makes PF more vulnerable to drought, leading to larger biomass losses and slower recovery. Their higher sensitivity to water disturbances results in lower resilience. In water-sufficient regions (humid areas), vegetation is limited by energy (e.g., temperature). PF are less sensitive to energy-related changes than NF. This reduced sensitivity helps PF maintain stability and explains their higher resilience in these humid conditions.

Figure S22. The inclination (e.g., slant vertical) of the letterings of forest management types below the figure should be improved for better readability.

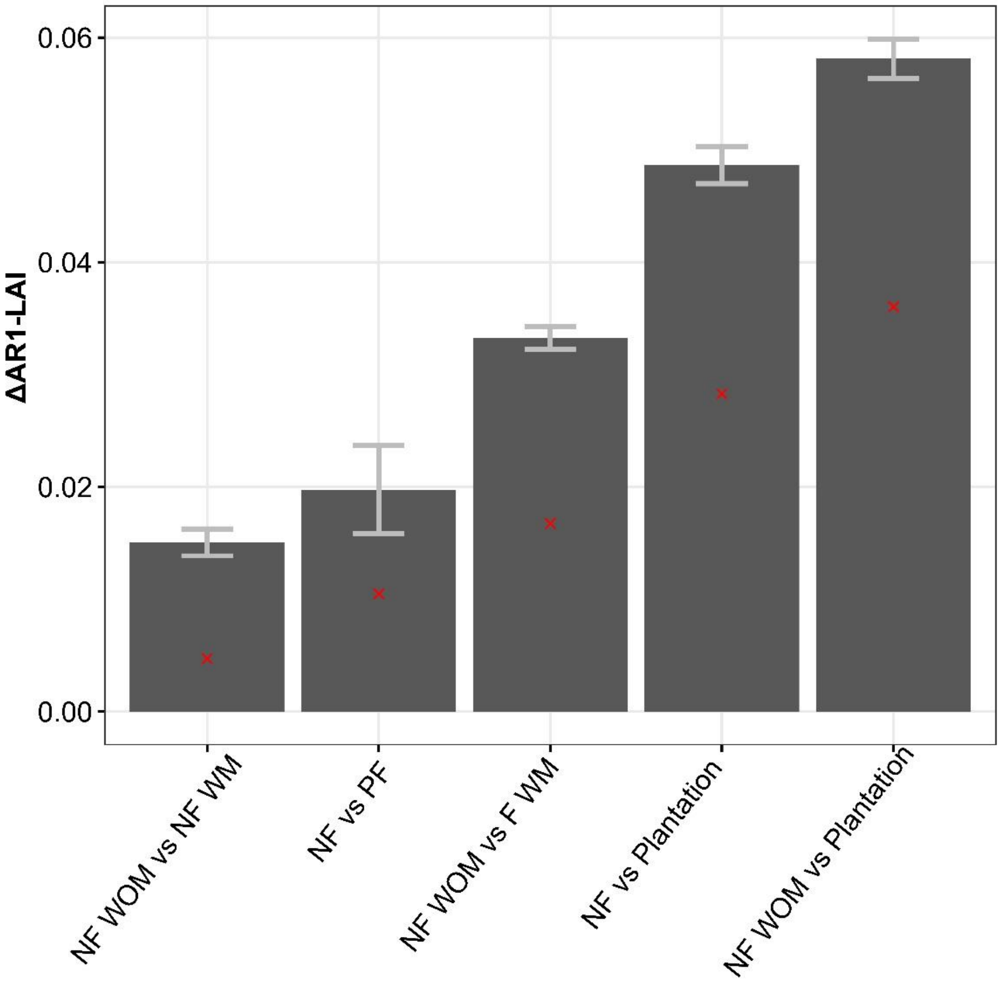
R: **Action taken**. We have corrected Figure in the revised version of the manuscript according to the reviewer's suggestion. Here is the modified figure:



Supplementary Figure S24. Resilience differences among forest management types estimated using λAR1 across multiple vegetation indices and three time-series pre-processing methods as described in Figure S22 caption (GS_Single, MVC, Single). Boxplots compare paired forest types (NF vs PF, NF vs Plantation, NF WOM vs F WM, NF WOM vs NF WM, NF WOM vs Plantation). Each point represents median values from a specific vegetation index and processing method (see legend). Crosses indicate non-significant differences, while squares enclosing crosses indicate significance relative to zero ($p < 0.05$). Red dashed line represents no resilience difference between two forest management types.

Figure S27. The labels of forest management types below the figure should be restricted to fall outside the figure.

R: **Action taken**. We have corrected Figure in the revised version of the manuscript according to the reviewer's suggestion. Here is the modified figure:



Supplementary Figure S32. Resilience differences (ΔAR1-LAI) among forest management types based on the independent crowdsourced dataset. Comparisons are performed using a nearest-neighbor matching method (see Supplementary Note 3 for details). Bars represent mean differences with error bars showing mean ± 2*SE, and red crosses indicate medians.

Does differences in species composition between forest ecosystems influenced the ecosystem resilience in this study? Was this factor considered as species richness, evenness and diversity under the vegetation variable, in addition to AGB variable? At least, using more vegetation variables in the model would help in detecting the actual influence of biophysical factors on forest ecosystem resilience.

R: In this study, we used vegetation variables such as aboveground biomass (AGB), belowground biomass (BGB), and root depth. However, we acknowledge that these metrics do not directly capture species compositional traits. Species composition—including richness, evenness, and diversity—can play a significant role in shaping forest ecosystem resilience.

Following the suggestions from the previous review round, we have in fact conducted preliminary explorations into the relationship between biodiversity and resilience (Fig. S31), which have confirmed the importance of this idea. Our initial findings indicate that regional biodiversity can significantly enhance ecosystem resilience, particularly for managed forests, where its influence is notably pronounced. However, it's important to clarify that despite this preliminary exploration, these heterogeneity data were not incorporated as direct input conditions for the models presented in this current analysis.

Therefore, we have now explicitly discussed this limitation in the revised manuscript's discussion section. We highlight that this analysis primarily focused on broad environmental variables (such as climate and soil) and acknowledge it as a preliminary, broad-scale investigation that did not deeply explore the direct influence of fine-scale biophysical factors like species diversity. This limitation has been clearly articulated in the revised manuscript.

In the future, we also plan to incorporate species composition variables—such as tree species richness, functional diversity, and canopy stratification—using recently available datasets (e.g., GEDI). Specifically, GEDI can accurately capture the three-dimensional spatial characteristics of vegetation, which will be invaluable for further exploring the intricate relationship between resilience and fine-scale vegetation features.

Action taken. We have corrected Figure in the revised version of the manuscript according to the reviewer's suggestion. Here is the modified text (Lines 509~513):

"Future research should focus on the integration of additional factors such as landforms and species diversity (including richness, evenness, and composition) into modeling to further improve our understanding of how management practices affect forest resilience under varying environmental conditions."

"This discrepancy is potentially due to the imbalance of samples in the crowdsourced points, where samples of forest managements like PF are less as compared to NF" – Please provide precise numbers in the manuscript text instead of using "less" and "more" qualifiers.

R: Action taken. We have corrected the Figure in the revised version of the manuscript according to the reviewer's suggestion. We added the number of samples in the Methods section:

"we collect ~210,000 crowdsourced validation samples ($N_{NF\ WOM}$ = 39534, $N_{NF\ WM}$ = 82358, $N_{PF\ without\ rotation}$ = 8357, $N_{PF\ with\ rotation}$ = 20247, $N_{Agroforestry}$ = 47751, and $N_{oil\ palms}$ = 11691) from crowdsourced datasets in 2015." (Lines 854~857)

Supplementary References:

1. Grimm, V. & Wissel, C. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia* **109**, 323–334 (1997).
2. Lenton, T. M., Livina, V. N., Dakos, V., van Nes, E. H. & Scheffer, M. Early warning of climate tipping points from critical slowing down: comparing methods to improve robustness. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **370**, 1185–1204 (2012).
3. Bathiany, S. *et al.* Ecosystem Resilience Monitoring and Early Warning Using Earth Observation Data: Challenges and Outlook. *Surv. Geophys.* **46**, 265–301 (2025).
4. Engineering, N. A. of. *Engineering Within Ecological Constraints*. (National Academies Press, 1996).
5. Arnoldi, J.-F., Loreau, M. & Haegeman, B. Resilience, reactivity and variability: A mathematical comparison of ecological stability measures. *J. Theor. Biol.* **389**, 47–59 (2016).
6. Scheffer, M. *et al.* Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
7. Scheffer, M. *et al.* Anticipating Critical Transitions. *Science* **338**, 344–348 (2012).
8. Verbesselt, J. *et al.* Remotely sensed resilience of tropical forests. *Nat. Clim. Change* **6**, 1028–1031 (2016).
9. Dakos, V. *et al.* Slowing down as an early warning signal for abrupt climate change. *Proc. Natl. Acad. Sci.* **105**, 14308–14312 (2008).
10. Forzieri, G., Dakos, V., McDowell, N. G., Ramdane, A. & Cescatti, A. Emerging signals of declining forest resilience under climate change. *Nature* **608**, 534–539 (2022).
11. Smith, T., Traxl, D. & Boers, N. Empirical evidence for recent global shifts in vegetation resilience. *Nat. Clim. Change* **12**, 477–484 (2022).
12. Boulton, C. A., Lenton, T. M. & Boers, N. Pronounced loss of Amazon rainforest resilience since the early 2000s. *Nat. Clim. Change* **12**, 271–278 (2022).
13. Liu, X., Wang, D., Chen, A. & Zeng, Z. Asymmetric sensitivity of boreal forest resilience to forest gain and loss. *Nat. Ecol. Evol.* **9**, 505–514 (2025).
14. Smith, T. *et al.* Reliability of resilience estimation based on multi-instrument time series. *Earth Syst. Dyn.* **14**, 173–183 (2023).
15. Smith, T. & Boers, N. Reliability of vegetation resilience estimates depends on biomass density. *Nat. Ecol. Evol.* **7**, 1799–1808 (2023).
16. Smith, T. & Boers, N. Global vegetation resilience linked to water availability and variability. *Nat. Commun.* **14**, 498 (2023).
17. Wang, Z. *et al.* Vegetation resilience does not increase consistently with greening in China's Loess Plateau. *Commun. Earth Environ.* **4**, 1–10 (2023).
18. Su, Y. *et al.* Pervasive but biome-dependent relationship between fragmentation and resilience in forests. *Nat. Ecol. Evol.* 1–15 (2025) doi:10.1038/s41559-025-02776-7.
19. Yao, Y. *et al.* Declined terrestrial ecosystem resilience. *Glob. Change Biol.* **30**, e17291 (2024).
20. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. & Walker, B. Catastrophic shifts in ecosystems. *Nature* **413**, 591–596 (2001).
21. Engineering, N. A. of. *Engineering Within Ecological Constraints*. (National Academies Press,

1996).

22. Woods, D. D. Four concepts for resilience and the implications for the future of resilience engineering. *Reliab. Eng. Syst. Saf.* **141**, 5–9 (2015).
23. Zhang, Y. *et al.* Spatial Heterogeneity of Vegetation Resilience Changes to Different Drought Types. *Earths Future* **11**, e2022EF003108 (2023).
24. Yang, X., Xu, X., Stovall, A., Chen, M. & Lee, J.-E. Recovery: Fast and Slow—Vegetation Response During the 2012–2016 California Drought. *J. Geophys. Res. Biogeosciences* **126**, e2020JG005976 (2021).
25. Buschke, F. & Brownlie, S. Reduced ecological resilience jeopardizes zero loss of biodiversity using the mitigation hierarchy. *Nat. Ecol. Evol.* **4**, 815–819 (2020).
26. Dakos V. *et al.* Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data. *PLOS ONE* **7**, e41010 (2012).
27. Kuehn, C. A mathematical framework for critical transitions: Bifurcations, fast–slow systems and stochastic dynamics. *Phys. Nonlinear Phenom.* **240**, 1020–1035 (2011).
28. Wissel, C. A universal law of the characteristic return time near thresholds. *Oecologia* **65**, 101–107 (1984).
29. Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. & Walker, B. Catastrophic shifts in ecosystems. *Nature* **413**, 591–596 (2001).
30. Nazarimehr, F., Jafari, S., Perc, M. & Sprott, J. C. Critical slowing down indicators. *Europhys. Lett.* **132**, 18001 (2020).
31. Zhang, Y. *et al.* Warming and disturbances affect Arctic-boreal vegetation resilience across northwestern North America. *Nat. Ecol. Evol.* **8**, 2265–2276 (2024).
32. Wingfield, M. J., Brockerhoff, E. G., Wingfield, B. D. & Slippers, B. Planted forest health: The need for a global strategy. *Science* **349**, 832–836 (2015).
33. Ciemer, C. *et al.* Higher resilience to climatic disturbances in tropical vegetation exposed to more variable rainfall. *Nat. Geosci.* **12**, 174–179 (2019).
34. van Nes, E. H. & Scheffer, M. Slow Recovery from Perturbations as a Generic Indicator of a Nearby Catastrophic Shift. *Am. Nat.* **169**, 738–747 (2007).
35. Rietkerk, M., Skiba, V., Weinans, E., Hébert, R. & Laepple, T. Ambiguity of early warning signals for climate tipping points. *Nat. Clim. Change* **15**, 479–488 (2025).
36. Yi, C. *et al.* Principles for guiding future research on resilience and tipping points. *Environ. Res. Lett.* **20**, 031008 (2025).
37. Moesinger, L. *et al.* The global long-term microwave Vegetation Optical Depth Climate Archive (VODCA). *Earth Syst. Sci. Data* **12**, 177–196 (2020).
38. Wulder, M. A. *et al.* Fifty years of Landsat science and impacts. *Remote Sens. Environ.* **280**, 113195 (2022).
39. Hemati, M., Hasanlou, M., Mahdianpari, M. & Mohammadimanesh, F. A Systematic Review of Landsat Data for Change Detection Applications: 50 Years of Monitoring the Earth. *Remote Sens.* **13**, 2869 (2021).
40. Xiao, J. *et al.* A review of remote sensing image spatiotemporal fusion: Challenges, applications and recent trends. *Remote Sens. Appl. Soc. Environ.* **32**, 101005 (2023).

41. Wang, Q. *et al.* A comprehensive review of spatial-temporal-spectral information reconstruction techniques. *Sci. Remote Sens.* **8**, 100102 (2023).
42. Lesiv, M. *et al.* Global forest management data for 2015 at a 100 m resolution. *Sci. Data* **9**, 199 (2022).
43. Sexton, J. O. *et al.* Global, 30-m resolution continuous fields of tree cover: Landsat-based rescaling of MODIS vegetation continuous fields with lidar-based estimates of error. *Int. J. Digit. Earth* **6**, 427–448 (2013).
44. Spawn, S. A., Sullivan, C. C., Lark, T. J. & Gibbs, H. K. Harmonized global maps of above and belowground biomass carbon density in the year 2010. *Sci. Data* **7**, 112 (2020).
45. Wang, Y. *et al.* Major forest increase on the Loess Plateau, China (2001–2016). *Land Degrad. Dev.* **29**, 4080–4091 (2018).
46. Yu, Z. *et al.* Natural forests exhibit higher carbon sequestration and lower water consumption than planted forests in China. *Glob. Change Biol.* **25**, 68–77 (2019).
47. Li, P., Li, H. & Si, B. Estimating deep soil water depletion and availability under planted forest on the Loess Plateau, China. *Sci. Total Environ.* **929**, 172228 (2024).
48. Duncan, B. N. *et al.* Space-Based Observations for Understanding Changes in the Arctic-Boreal Zone. *Rev. Geophys.* **58**, e2019RG000652 (2020).
49. Zhang, Y., Keenan, T. F. & Zhou, S. Exacerbated drought impacts on global ecosystems due to structural overshoot. *Nat. Ecol. Evol.* **5**, 1490–1498 (2021).
50. Chen, M. *et al.* Response of soil water to long-term revegetation, topography, and precipitation on the Chinese Loess Plateau. *CATENA* **236**, 107711 (2024).
51. Yu, Z. *et al.* Natural forest growth and human induced ecosystem disturbance influence water yield in forests. *Commun. Earth Environ.* **3**, 1–8 (2022).
52. Wang, C., Zhang, W., Li, X. & Wu, J. A global meta-analysis of the impacts of tree plantations on biodiversity. *Glob. Ecol. Biogeogr.* **31**, 576–587 (2022).
53. Abdelmoaty, H. M. & Papalexiou, S. M. Changes of Extreme Precipitation in CMIP6 Projections: Should We Use Stationary or Nonstationary Models? (2023) doi:10.1175/JCLI-D-22-0467.1.
54. Hashimoto, H. *et al.* Exploring Simple Algorithms for Estimating Gross Primary Production in Forested Areas from Satellite Data. *Remote Sens.* **4**, 303–326 (2012).
55. Zong, X., Liu, Y. & Yin, Y. Identifying the dominant compound events and their impacts on vegetation growth in China. *Weather Clim. Extrem.* **45**, 100715 (2024).
56. Deng, Y. *et al.* Responses of vegetation greenness and carbon cycle to extreme droughts in China. *Agric. For. Meteorol.* **298–299**, 108307 (2021).

Response to Reviewers

The authors sincerely thank the anonymous reviewers for their valuable and constructive feedback. We have carefully addressed all concerns and revised the manuscript accordingly. Below, we provide detailed responses to each comment from reviewers. In the revised manuscript, reviewer comments are shown in black, our revisions are highlighted in red, and our responses are indicated in blue. Please note that all line numbers refer to the revised version of the manuscript. Additionally, the references cited within the red-colored text are available in the tracked manuscript, while those in the blue responses are included in this document.

Response to Reviewer #2

Reviewer #2 (Remarks to the Author):

A lot of effort has gone in the rebuttal letter and in improving the robustness of the analyses and the main text. However, I still find that some issues were not fully answered and some clarifications still need to be included. Below my main concerns and some minor comments.

R: We sincerely thank the reviewer for your positive assessmen of our previous revisions and for the valuable time and effort devoted to providing detailed feedback. Following your comments, we have carefully addressed all remaining issues and incorporated the requested clarifications and analyses into the revised manuscript and Supplementary Information.

1.The definition of "resilience" seems to be a key point in your response, however, I found the definitions given to be unclear. Out of the sources you mentioned at the start of your response, as far as I can tell, four do not define resilience (Dakos et al., 2008; Lenton et al., 2012; Scheffer et al., 2009, 2012), three define resilience as variations of "return to the reference state after a disturbance" (Arnoldi et al., 2016; Grimm & Wissel, 1997; Verbesselt et al., 2016), i.e. what you define as engineering resilience, and only the remaining two references differentiate between engineering vs. ecological resilience, defining ecological resilience as the "magnitude of disturbance that can be absorbed before the system changes [to a different state]" (Bathiany et al., 2025; Holling, 1996). The subsequent sources you cite later in your response also only give mixed definitions that include both the ability to resist disturbances as well as the ability to recover from disturbances (Buschke & Brownlie, 2020; Forzieri et al., 2022; Smith et al., 2022). These two concepts (resist vs recover) should be differentiated more clearly. The way you describe ecological resilience as "concerning the system's continuous response to a sequence of disturbances, both strong and weak" does not seem to reflect the cited references. Later in your response, you also mention that the indicators you are using "are also closely related to the recovery rate after disturbance", making the whole concept even more unclear: if you are measuring the recovery rate after a disturbance, you seem to relate the study more to the engineering resilience, contrary to what you state.

Aside from the somewhat confusing rebuttal letter, in the main text you define resilience as: "the capacity of ecosystems to withstand and recover from external perturbations while preserving their equilibrium state" (L72-74). This seems to be a mix between engineering ("recover") and ecological ("withstand") resilience, while the cited source differentiates between "the rate of recovery upon perturbation" (engineering resilience) and "the size of alternative attraction basins" (ecological resilience) (Scheffer et al., 2015). Finally, your definitions of resilience in the supplementary materials ("the ability of a system to persist and maintain its original state in the face of a continuous sequence of both strong and weak disturbance") does not reflect the cited source, which only describes resilience as "the return rate from perturbations" (Boulton et al., 2022). At no point the cited references mention the concept of "continuous disturbance" or differentiate between "strong and weak disturbance". Since the terminology around these concepts can be very confusing and has been used with different meanings in different contexts, it is important to be very clear about your definitions, keep them consistent across the whole text, and cite the correct sources. If I understand correctly, what you describe as "ecological resilience" could be termed "resistance" or "tolerance" as defined by Grimm & Wissel (1997), which is also in line with a more recent overview of the terminology by Van Meerbeek et al. (2021) (also cf. "latitude" as the "distance to threshold on the system variable axis"). These terms (e.g., resistance) might more intuitively convey what you are describing. You also use the term "recovery" almost interchangeably with "resilience", but these two terms should be differentiated more clearly.

Whatever terminology you choose to use be sure to define it and cite it correctly to avoid confusion, and keep it consistent throughout the text.

R: We sincerely thank the reviewer for the careful and constructive comments regarding the definition of "resilience." We fully agree that the distinction between engineering resilience and ecological resilience must be made explicit and used consistently throughout the manuscript.

Action Taken. In the revised version, we have carefully unified the terminology and clarified the definition of ecological resilience in both the main text and the Supplementary Materials. This definition follows Holling (1996) and Bathiany et al. (2025), who characterize ecological resilience as the [*magnitude of disturbance that can be absorbed before the system shifts to an alternative state]*, and is conceptually aligned with Grimm & Wissel (1997) and Van Meerbeek et al. (2021), who define this resilience with [*system resistance* or *tolerance disturbances]*. Specifically, based on these references recommended by the reviewer (e.g., Holling, 1996; Bathiany et al., 2025; Van Meerbeek et al., 2021; Grimm & Wissel 1997)[1–5], we ensure conceptual accuracy and completeness in the texts. We now define ecological resilience as: "the capacity of a system to resist (or absorb) changes from disturbances, maintain its present structure and functions". We also cited these four papers in the text. Here are the modified texts:

"Forest management influences key ecosystem services essential for societal well-being. However, how these human activities affect forest resilience (i.e., the system's capacity to resist changes and continue to function in its present state) under different environmental conditions remains poorly understood." (Lines 5~8)

"To address these complexities, we adopt the concept of ecological resilience, defined as the capacity of a system to resist (or absorb) changes from disturbances, maintain its present structure and functions[33–36]. In other words, ecological resilience reflects the ability to tolerate perturbations without shifting to a qualitatively different system." (Lines 71~75)

“In contrast, ecological resilience refers to a system’s capacity to resist (or absorb) changes, maintain its present structure even under a series of disturbances of varying intensity[39,43–45].” (Supplementary Note 9, lines 297~299)

In addition, we have revised terms such as "recovery" and "recover" in the entire text to avoid ambiguity of ecological resilience. We now directly use the term “resilience” or expressions such as “the ability to absorb disturbances (Holling, 1996; Bathiany et al., 2025)” in the texts. Here are some related modified sentences:

“Consequently, managed forests can be exposed to greater water and heat stress, which impairs their ability to absorb disturbances (i.e., lower resilience)[44].” (Lines 386~387)

“Deep-rooted forests in NF WOM can buffer drought impacts and maintain its essential function by accessing groundwater[48], leading to higher resilience.” (Lines 426~428)

2. Definition of tipping points: you explain very in detail what tipping points are from a theoretical perspective. The addition of supplementary note 8 will definitely help clarify these notions to the readers. However, reviewer #1 asked specifically to define tipping points for your PF and NF. You only provide a general example where a forest turns into shrubland/grassland following severe a repeated disturbance. This does not differentiate between NF and PF and, importantly, it does not seem to align with the constraints set in your study (only forest areas). Within your study framework, are you seeing any such state shifts or tipping point? Is your filtering methodology intrinsically excluding any tipping point altogether?

R: We sincerely thank the reviewer for this valuable comment. We agree that the previous example (forest transitioning into shrubland or grassland) was not fully consistent with the scope of our study, which is restricted to forested areas. In the revised version, we have provided a more relevant example: a transition from high-density closed-canopy forests to low-density forests, which still reflects structural and functional changes within forest ecosystems (i.e., tipping point). Furthermore, we have refined this example to better align with different forest management types (PF and NF) (see Action Taken).

Moreover, the focus of our study is not to directly identify or monitor tipping points, but rather to assess differences and dynamics in forest resilience. Similar to Forzieri et al. (2022), our analysis is limited to forest ecosystems, using early warning indicators (EWI) to infer the capacity of forests to absorb disturbances, maintain their core structure and functions.

We would also like to clarify the conceptual link between indicators and tipping points. For indicators (e.g., AR1) to be interpreted as early warnings of critical transitions, there must first be independent evidence that the system can exhibit such tipping behavior, that is, transitioning from one stable state to another. In other words, the increase in indicators such as AR1 (i.e., a loss of resilience) does not necessarily imply that the system is approaching a tipping point; additional evidence is required to demonstrate that the system is indeed capable of experiencing such a transition. Only under this condition can the increase in indicators such as AR1 be interpreted as an elevated risk of tipping point occurrence. Previous studies suggest that regions such as the Amazon[6–8], temperate forests[9], and boreal forests[10,11] have already exhibited signs of degradation or structural change under disturbances such as fire and drought[12]. These regions are expected to experience further large-scale transformations in the future due to ongoing warming[12–14].

In systems without such transitions, indicators (e.g., AR1) can be used to monitor resilience changes. The physical meaning of indicators inherently reflects the system's resilience. For instance, AR1 represents temporal memory effects, indicating the extent to which disturbances in previous time steps persist in subsequent periods. Accordingly, in Forzieri's (2022) study, they concluded that global forests are showing a decline in resilience rather than an increase in tipping point risk. Similar studies have also primarily focused on resilience rather than the direct identification of tipping points[10,15,16]. Our study follows a similar rationale, comparing long-term resilience differences between forests under distinct management conditions rather than identifying tipping events.

Most of the analyses in this study are based on the long-term resilience (Fig. 1~5), with only Figure 6 involving the dynamic changes in resilience. The majority of the perspectives presented in this paper rely on using these CSD indicators as proxies for resilience. In the Supplementary Materials (Supplementary Notes 7~8), we use sections to introducing the CSD framework and its associated indicators (e.g., AR1), given their intrinsic theoretical and historical connections. The CSD framework provides the physical and mathematical foundation for these indicators (e.g., AR1), which has driven their development and application in resilience studies. Moreover, introducing the CSD framework helps readers better understand the ecological meaning of these indicators. For example, interpreting ecosystem resilience from the perspective of the basin of attraction can offer theoretical support for our subsequent interpretation of forest resilience.

Action Taken. We have provided a more relevant example (related with the PF and NF) for tipping point:

"In the context of our study, the tipping point is defined as a theoretical threshold at which the high-density closed-canopy forests would lose their resilience and potentially shift toward an alternative state such as low-density forests[48]. Although PF and NF currently exhibit similar states in photosynthesis, PF generally possess lower structural complexity and shallower, simplified root systems. These characteristics reduce their ability to absorb external disturbances, making them more prone to crossing the tipping point and transitioning toward low-cover, open forests under continued disturbances. In contrast, NF, with their higher structural complexity and well-developed root systems, tend to maintain their initial high-cover, closed-canopy state due to stronger self-regulation and resilience." (Supplementary Note 9, Lines 315~323)

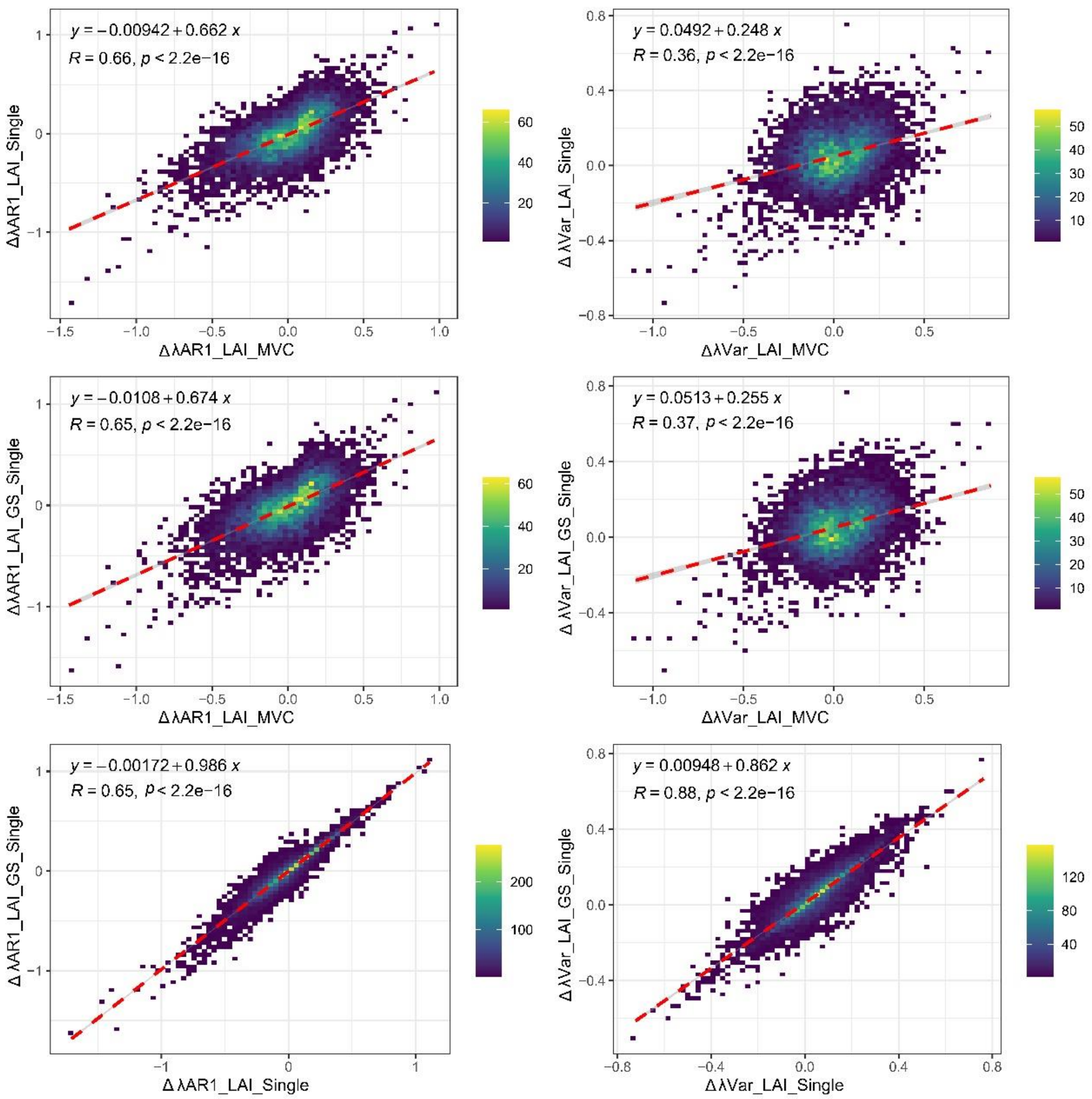
In addition, we have added a discussion in the Supplementary Materials about the relationship between resilience indicators and tipping points.

"Early warning indicators (EWI, e.g., λAR1 and λVar) are to be interpreted as a measure of an approaching tipping point[30,31]. In systems that lack such critical transition, a decline in resilience does not necessarily imply an approaching tipping point. This study focuses on global forest ecosystems. Previous studies reported that some regions such as the Amazon, temperate forests, and boreal forests have exhibited signs of degradation or structural transition under disturbances such as fire and drought[32–35]. These regions are expected to experience further large-scale transformations in the future due to ongoing warming[36,37]. Confirming the widespread occurrence of critical transitions in global forests would strengthen the case for using EWI as a measure of proximity to tipping points.

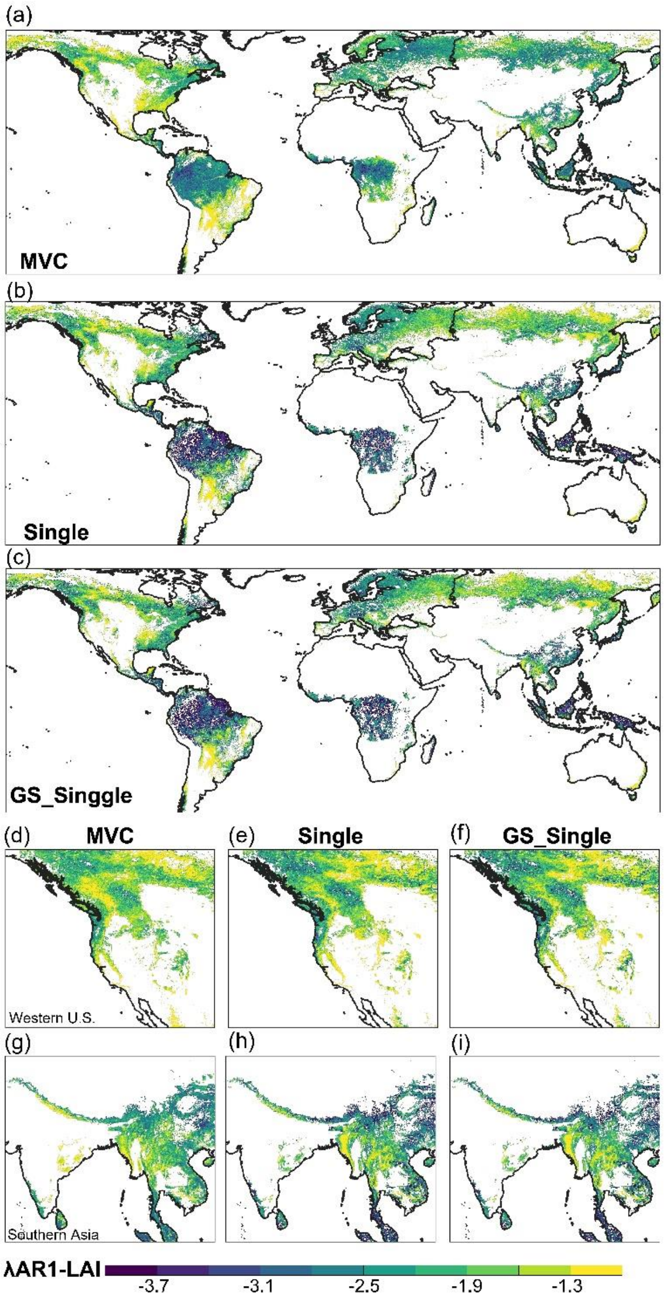
In systems without tipping points, elevated EWI values can still indicate a loss of system resilience. EWI have physical interpretations related to system resilience. For example, an increase in AR1 reflects stronger memory of past states, meaning that previous disturbances continue to influence the present[38]. This suggests a reduced ability of the system to absorb disturbances and maintain its equilibrium state. In this study, we used the EWI to explain resilience difference across managed forests." (Supplementary Note 8, Lines 259~273)

3-1. Temporal resolution of indices: it is good that you included sensitivity tests. However, the provided figures do not give enough detail to fully evaluate these tests. E.g., Fig. S22 does not provide explained variance and goodness of fit metrics, and the colorbar is not described; Fig. S25 only show two regions (where I would argue that some differences can already be seen): what about other areas? Why show only these two?

R: Thank you for this important comment. **Action taken.** We have corrected the figures in the revised version of the manuscript according to the reviewer's suggestion. Specifically, in Fig. S22, we have added the explained variance (R) and *p-value* within each panel and explicitly described the color bar as representing the density of spatial samples (darker colors = lower density; brighter colors = higher density). In addition, Fig. S25 has been expanded to include representative regions (Southeastern U.S., Amazon and Southeast Asia) and globe to provide a more comprehensive spatial comparison.



Supplementary Figure S22. The impacts of time-series pre-processing methods on the resilience difference between PF and NF (e.g., ΔλAR1-LAI$_{(PF,NF)}$). There are three time-series pre-processing methods for estimating resilience by: (1) monthly vegetation index series aggregated via maximum value compositing from MODIS Terra and Aqua LAI datasets (i.e., ΔλAR1_LAI_MVC); (2) the MODIS Terra LAI datasets with its original temporal resolution (no temporal aggregation; i.e., ΔλAR1_LAI_Single); and (3) the Terra time series restricted to the growing season at its original 8-day resolution (i.e., ΔλAR1_LAI_GS_Single). Each scatterplot includes a fitted linear regression (red dashed line) with the regression equation, coefficient of determination (*R*), and significance level (*p*) shown in the panel. The color bar represents the density of spatial samples, with darker colors indicating lower sample density and brighter colors representing higher density (i.e., more grid cells per bin). The strong linear relationships indicate that the choice of pre-processing method has only minor effects on resilience estimates.



Supplementary Figure S25. Spatial patterns of long-term λAR1-LAI are shown for three regions (globe, the western U.S., and southern Asia) to illustrate the effect of time-series pre-processing (as described in Figure S22 caption including GS_Single, MVC, and Single) on the resilience estimation. There are three time-series pre-processing methods for estimating resilience by: (1) monthly vegetation index series aggregated via maximum value compositing from MODIS Terra and Aqua LAI datasets (i.e., ΔλAR1_LAI_MVC); (2) the MODIS Terra LAI datasets with its original temporal resolution (no temporal aggregation; i.e., ΔλAR1_LAI_Single); and (3) the Terra time series restricted to the growing season at its original 8-day resolution (i.e., ΔλAR1_LAI_GS_Single).

3-2. Moreover, you tested if the early warning indicators can detect changes in forest “resilience” prior to disturbance events (Fig. R4). Here, it is not clear 1) how you identified the two disturbance events; 2) how you tested for "significant" increases prior to the disturbance events; and 3) how you can state that "in both periods, the early warning indicators […] showed significant increases prior to the disturbances", while there is no calculated value for the indicators prior to 2003 (and thus prior to the first disturbance event shown in Fig. R4).

R: We appreciate the reviewer's detailed comments and highlight that the interpretation of Figure R4 may not have been sufficiently clear. The reviewer's mainly concern relate to (1) how disturbance events were identified, (2) how "significant" increases were tested, and (3) how the indicator values prior to 2003 were determined. We provide the following clarifications.

(1) Identification of disturbance events.

Disturbance events were identified directly from abrupt negative anomalies in the vegetation index (VI) time series, where VI values dropped more than one standard deviation below the long-term mean. Such abrupt deviations have been widely used in previous studies as proxies for disturbance signals. Two distinct periods (2001–2005 and 2012–2015) were identified as major disturbance intervals based on these deviations.

(2) Definition of "significant" increases.

The term "significant increase" was not used appropriately in this context, as it was originally intended to describe a rapid rise rather than a statistically tested change. Here we rephrased it to "substantial increase".

(3) Computation and interpretation of the AR1 time series.

The dynamics of AR1 indicators were calculated using a 4-year moving window over the 2001–2015 record. Therefore, the AR1 value for 2003 corresponds to the 2001–2005 window, and that for 2013 corresponds to the 2011–2015 window.

Finally, we would like to return to the reviewer's earlier question (related to the Figure R4) in the previous round — *["Are changes in greenness over 5×5 km detectable with a frequency of one or two observations per month, and can they truly capture non-abrupt, non-anthropogenic disturbances?" by Reviewer 1]*

To address it, we presented Figure R4 to show that deviations in the original vegetation time series are indeed detectable and correspond to periods of anomalous system behavior (i.e., response to disturbances).

Moreover, we have performed a series of sensitivity analyses to test the effects of different spatial (500 m vs. 1 km, Fig. S29) and temporal (8-day vs. monthly, Figs. S22-S25 and S30) resolutions on resilience estimation. The results show minimal differences, indicating that our findings are robust to data resolution. These pieces of evidence collectively demonstrate that the influence of spatial and temporal resolution on our results is minimal, and that the approach can detect relevant disturbances, as reflected by the deviations of the vegetation index (VI) from its mean state.

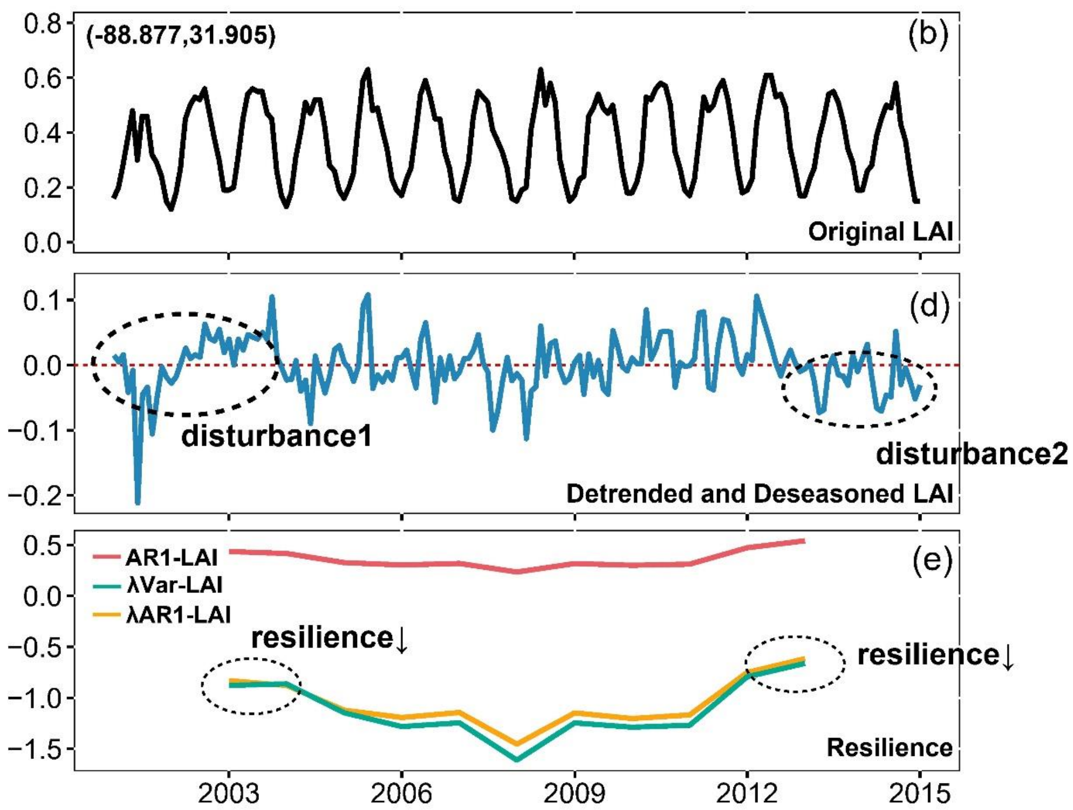


Figure R4 of Reversion 2. Figure from Figure1 in the main text. (b) Original LAI time series at the given location, showing strong seasonal patterns. (d) Detrended and deseasonalized LAI, with two major disturbance events (disturbance1 and disturbance2) indicated by sudden declines. (e) Early warning indicators (AR1-LAI, λVar-LAI, and λAR1-LAI) calculated from the processed LAI time series. Notably, all indicators show significant increases (i.e., resilience decline) preceding and during the disturbance periods, demonstrating the sensitivity of CSD-based metrics to disturbance-induced changes in ecosystem resilience.

3-3. On this note, in the main text you have to be careful about this distinction. E.g. at line 112-113 you state that "[CSD] indicators are computed over the entire observational period", after defining the observational period as 2001–2015. From my understanding, this is not true, since your 4-year moving window approach cuts the start and end of the timeseries. Please clarify.

R: We appreciate the reviewer's detailed comments. Two distinct calculations were performed in our study. Please note that the dynamic changes in CSD indicators are only related to Fig. 6, whereas Figs. 1–5 are based on the long-term mean early-warning indicators. The "long-term resilience" (Figs. 1–5) is derived once from the full 2001–2015 vegetation series to characterize the average state of each forest pixel. The "temporal dynamics of resilience" (Fig. 6) are based on 4-year moving windows (2003–2013). The long-term resilience was not derived from averaging the moving-window results. This approach enables the detection of evolving trends in resilience without influencing the long-term estimates.

As described in the Methods section, two distinct calculations of CSD indicators are as follows:

"Forest resilience is quantified as a long-term ecological characteristic and derived from the CSD indicators computed over the whole time series (2001–2015) at each pixel (1-km) (e.g., 2001–2015 λAR1-LAI). This estimate describes the long-term average resilience of forest ecosystems. We complement this analysis by exploring the temporal evolution of forest resilience, deriving CSD indicators over 4-year moving windows, as proposed in recent studies[42]. This results in a time series of resilience values over the period 2003–2013 at each pixel. Resilience dynamics are then expressed as linear trend of the time-varying indicators (e.g., λAR1-LAI trend)." (Lines 667~674)

Here are the processing methods:

(1) The monthly vegetation index (VI) data for 2001–2015 were first detrended and deseasonalized to remove seasonal cycles and long-term trends;
(2) Based on the processed VI data (i.e., 2001-2015), we calculated the long-term CSD indicators (e.g., 2001–2015 λAR1–LAI) to represent the average resilience of forest ecosystems;
(3) To capture temporal dynamics in resilience, we further applied a 4-year moving window to the 2001–2015 record to compute λAR1–LAI for each window and then derived the corresponding trend (e.g., λAR1–LAI trend).

3-4. Finally, Fig. R5 is supposed to show increases in the indicators prior to a major disturbance event. Yet, the indicators show an increase throughout the whole period in which they were calculated, so how can this increase be related only to the disturbance event and rather than a longer-term trend?

R: Thank you for this important comment. We fully agree that the continuous increase in AR1 throughout the analyzed period may reflect a longer-term trend. The analysis shown in Figure R5 was intended as a qualitative illustration, using published examples to demonstrate that vegetation responses to disturbances can be detected at a 16-day temporal and 1-km spatial resolution. This Figure R5 specifically responds to the reviewer's initial question: *["Are changes in greenness over 5×5 km detectable with a frequency of one or two observations per month, and can they truly capture non-abrupt, non-anthropogenic disturbances?" by Reviewer 1]*

Nevertheless, our interpretation also has basis. Previous studies have reported that abrupt disturbances such as wildfires and droughts can trigger short-term and pronounced increases in early-warning indicators (e.g., AR1) immediately before the events[10,17,18]. These findings support our qualitative observation that the rapid increase in AR1 shown in Figure R5 may be associated with disturbances. Based on reviewer's comment and previous studies, we would also argue that distinguishing the effects of long-term climate trends from short-term extreme events on resilience change is a valuable direction for future research.

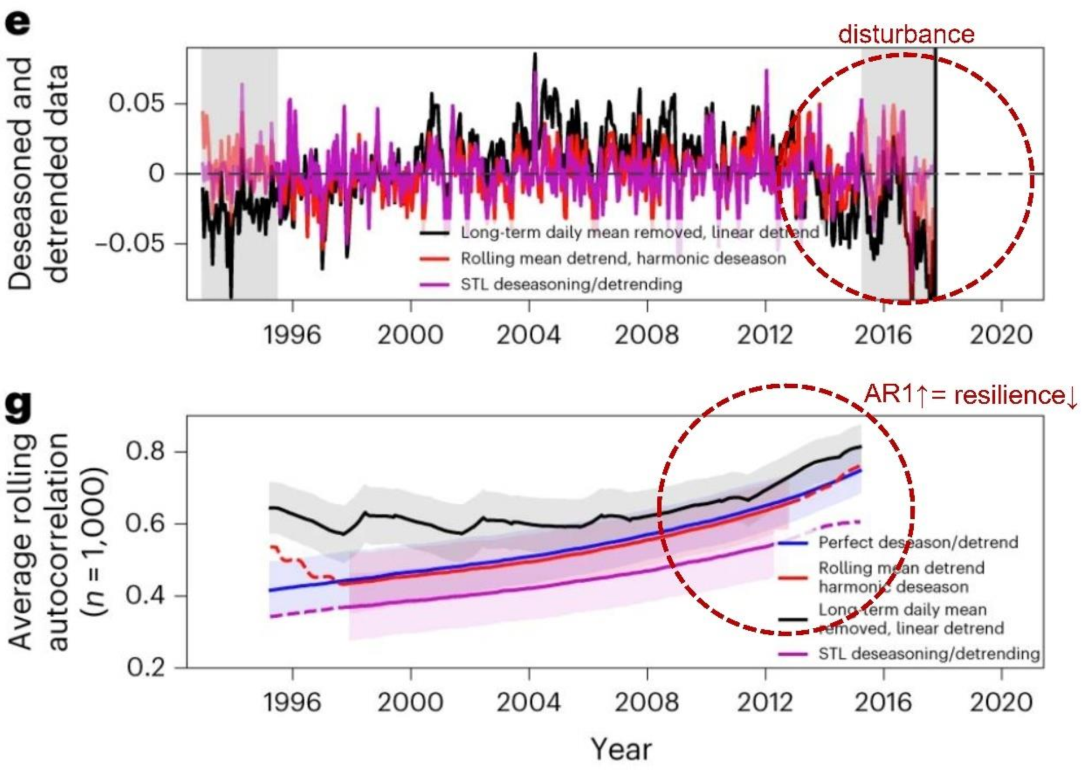


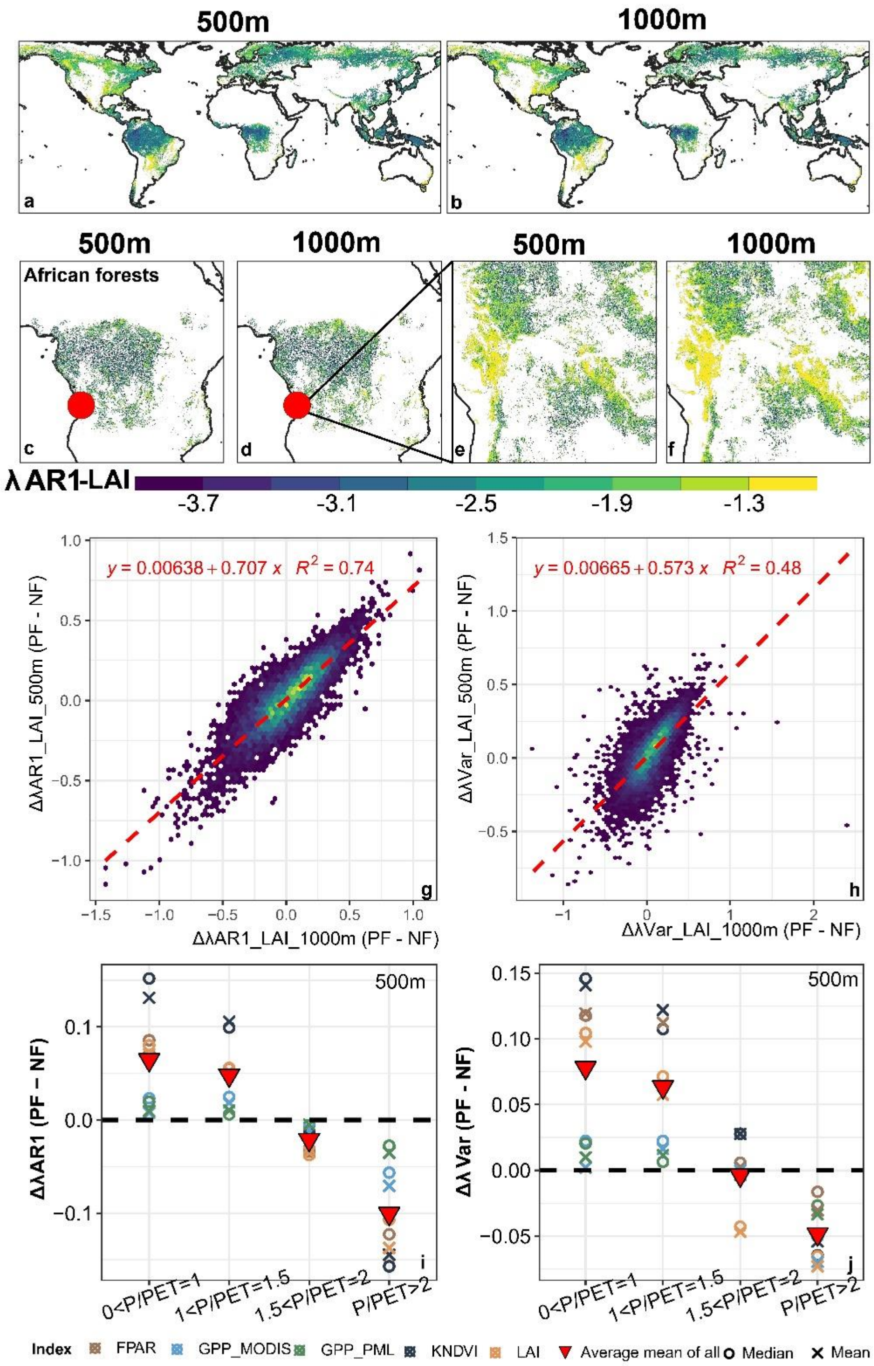
Figure R5 of Reversion 2. Detrending and deseasoning of NDVI time series and corresponding rolling AR1 calculations. (e) NDVI time series (16-day resolution) after removing seasonality and trends using different methods: long-term daily mean removal with linear detrending (black), rolling mean detrending with harmonic deseasoning (red), and STL-based deseasoning/detrending (magenta). (g) Average rolling autocorrelation (AR1) calculated over a moving window (n = 1,000) for each preprocessing method. The comparison shows that different deseasoning and detrending approaches can impact the estimated AR1 values, highlighting the importance of preprocessing choices when applying early warning indicators to remote sensing data. This figure is modified from ref[15].

4. Spatial resolution: reviewer #1 seemingly misunderstood this value as, clearly, 5 km ≠ 1 km, and your sensitivity analysis on the spatial resolution lends further robustness to your choice. Still, one nitpick in line with previous comments is that your provided evidence in Figure S29 is only based on qualitative visual analysis of the maps over African forests. Can you be sure that these two maps and the underlying data are statistically indistinguishable? And why are you only showing Africa? A suspicious reader might think that you are only showing what you want to show.

R: A direct comparison between the 500 m and 1000 m raster maps of resilience would require aggregating the 500 m data to match the 1000 m resolution, which could violate the premise of the sensitivity analysis and introduce artificial spatial smoothing.

In addition to the qualitative comparison presented in maps in Figure S29, we therefore conducted a quantitative statistical assessment to verify that the resilience metrics are consistent across different spatial resolutions. Specifically, we used the global 500 m resolution resilience index and the forest-management dataset to perform a moving-window analysis, comparing the resilience differences between PF and NF. Within paired windows, we compared the resilience differences derived from the 500 m and 1000 m datasets and found them to be highly consistent (Fig. S29e, f). Moreover, based on the 500 m data, we observed that $\Delta\lambda AR1_{(PF, NF)}$ ($\Delta\lambda Var_{(PF, NF)}$) gradually decreases with increasing P/PET (Fig. S29g, h), which aligns with the results obtained from the 1000 m analysis in the main text (Fig. 3a-c).

Action Taken. Based on the reviewers' comments, we have added a global map here.



Supplementary Figure S29. Effect of spatial resolution on forest resilience estimates at 500 m and

1000 m. (a–b) Global maps of λAR1–LAI at 500 m and 1000 m spatial resolutions. (c–f) Comparison of λAR1–LAI patterns in African tropical forests at both resolutions, showing consistent large-scale patterns with localized differences. (g–h) Relationships of resilience differences (ΔλAR1–LAI and ΔλVar–LAI between PF and NF) derived at 500 m versus 1000 m, with corresponding regression lines (red dashed) and coefficients of determination (R^2). (i–j) Mean resilience differences (PF − NF) along precipitation-to-potential-evapotranspiration (P/PET) gradients at 500 m resolution. Each marker represents a vegetation index (FPAR, GPP_MODIS, GPP_PML, kNDVI, and LAI), and red triangles denote the mean across indices.

5. Fig. S21 and others do not provide a sufficient explanation about what the y axis is showing and what statistical test was used to test the differences between management types. The reader has to refer to Figure 2, which makes the reading experience unnecessarily convoluted. Please add all relevant information in all figures' descriptions.

R: **Action Taken.** We have revised the figure caption to explicitly clarify the meaning of the y-axis and the statistical test used. The revised caption now reads as follows:

"**Supplementary Figure S21.** Sensitivity test of different filtering thresholds in paired grid method. The y-axis represents the difference in resilience (ΔAR1) between two forest management types (e.g., NF vs PF), where a positive ΔAR1 indicates higher resilience (i.e., lower AR1) in the first forest type (NF). Asterisks denote statistical significance of ΔAR1 differences between paired management types (*$p < 0.05$; N.S., not significant), assessed using the Wilcoxon test. Each boxplot summarizes ΔAR1 across five vegetation indices (FPAR, GPP_MODIS, GPP_PML, kNDVI, and LAI) based on all paired grids that satisfy the corresponding environmental similarity constraints. Panels (a–g) progressively apply stricter thresholds on the differences in elevation (ΔDEM), forest age (ΔAge), forest cover change (ΔFCC), and aboveground biomass (ΔAGB): (a) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 10%; (b) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 5%; (c) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 3%; (d) ΔDEM < 500 m, ΔAge < 10 yr, ΔAGB < 10 Mg ha^{-1}; (e) ΔDEM < 500 m, ΔAge < 10 yr, ΔAGB < 5 Mg ha^{-1}; (f) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 10% and ΔAGB < 10 Mg ha^{-1}; (g) ΔDEM < 500 m, ΔAge < 10 yr, ΔFCC < 5% and ΔAGB < 5 Mg ha^{-1}. Each group of boxplots compares resilience differences between management types, including NF versus PF, NF versus Plantation, NF WOM versus F WM, NF WOM versus NF WM, and NF WOM versus Plantation." (Supplementary, Lines 578~593)

Following your comments, we have revised the captions of all relevant supplementary figures (including Figs. S14, S17, S18, S19, S21, and S27) to provide complete and self-contained descriptions. We intentionally retained the phrase "Similar to…" at the beginning of each caption to help readers easily locate and compare corresponding sensitivity analyses in the main text. In addition, we have added concise explanations of the key elements in each figure, including the definition of the y-axis, the statistical tests used, and other essential details to ensure that each figure can be understood independently. Here are modified captions in the supplementary figures:

"**Supplementary Figure S17**. Impact of biophysical factors on resilience difference between different forest management types. Similar to Fig. 4 (and Figure S9b, d), but for the resilience differences between different forest management types. The "Location_Factor" indicates the local biophysical contexts of grids (y-axis). SHAP summary plots showing the relative importance of environmental and site factors in predicting the resilience differences across forest management types. Each point represents one sample: colors indicate the feature value (red = high, blue = low), while the SHAP value shows the strength and direction of the feature's effect on the model output (positive = increasing, negative = decreasing). For example, red points with positive SHAP values indicate that high values of the predictor have a positive effect on the target variable, while blue points with negative SHAP values suggest that low predictor values contribute negatively. Four comparisons are presented: (a) NF vs. Plantations, (b) unmanaged natural forests (NF WOM) vs. managed forests (F WM), (c) NF WOM vs. managed NF (NF WM), and (d) NF WOM vs. Plantations. The R^2 and *KGE* in the lower right corner are derived from the 5-fold cross-validation." (Supplementary, Lines 539~551)

"**Supplementary Figure S19**. Sensitivity test of resilience differences between forest management types to forest age thresholds in paired-grid selection. Similar to Fig. 2k, but with different thresholds of forests age difference (5 and 7 years) in paired-grid method. Boxplots show resilience differences among forest management types (ΔλAR1, y-axis) under alternative thresholds of forest age difference: < 5 years (top) and < 7 years (bottom), compared with the 10-year threshold used in Fig. 2. Results are presented for five vegetation indices (FPAR, GPP_MODIS, GPP_PML, kNDVI, and LAI). Asterisks denote significant differences ($p < 0.05$), while N.S. indicates non-significance, assessed using the Wilcoxon test." (Supplementary, Lines 563~570)

6. Standardization: in supplementary note 4 you mention that "Standardization of vegetation index data could influence resilience metrics. It is more likely to affect λVar, because standardization sets the standard deviation of the data to one.". That being said, you only tested the effect of standardization on λAR1. Here the appropriate sensitivity test would be to test it on λVar instead.

R: Because standardization changes the variance structure of the data and can directly affect the estimation of λVar, we did not standardize the vegetation index time series in our main analysis.

Action Taken. Although standardization was not applied in the main workflow, we conducted an additional sensitivity test in response to Reviewer #1's and your comment to examine whether standardization could influence the results. Specifically, we compared two preprocessing schemes for kNDVI: (1) Anomaly series, obtained by removing the long-term mean (detrended deviations from the mean); and (2) Z-score standardized series, obtained by subtracting the long-term monthly mean and dividing by the corresponding monthly standard deviation.

As shown in Figure R1, both λAR1 and λVar patterns remain highly consistent between the anomaly-based and Z-score-based approaches. Spatially, the resilience maps exhibit similar global distributions (Fig. 1Ra-d). The resilience difference between PF and NF show negative relationship with the P/PET across two preprocessing methods (Fig. 1Rg-j). This confirms the robustness of our results with respect to vegetation index standardization.

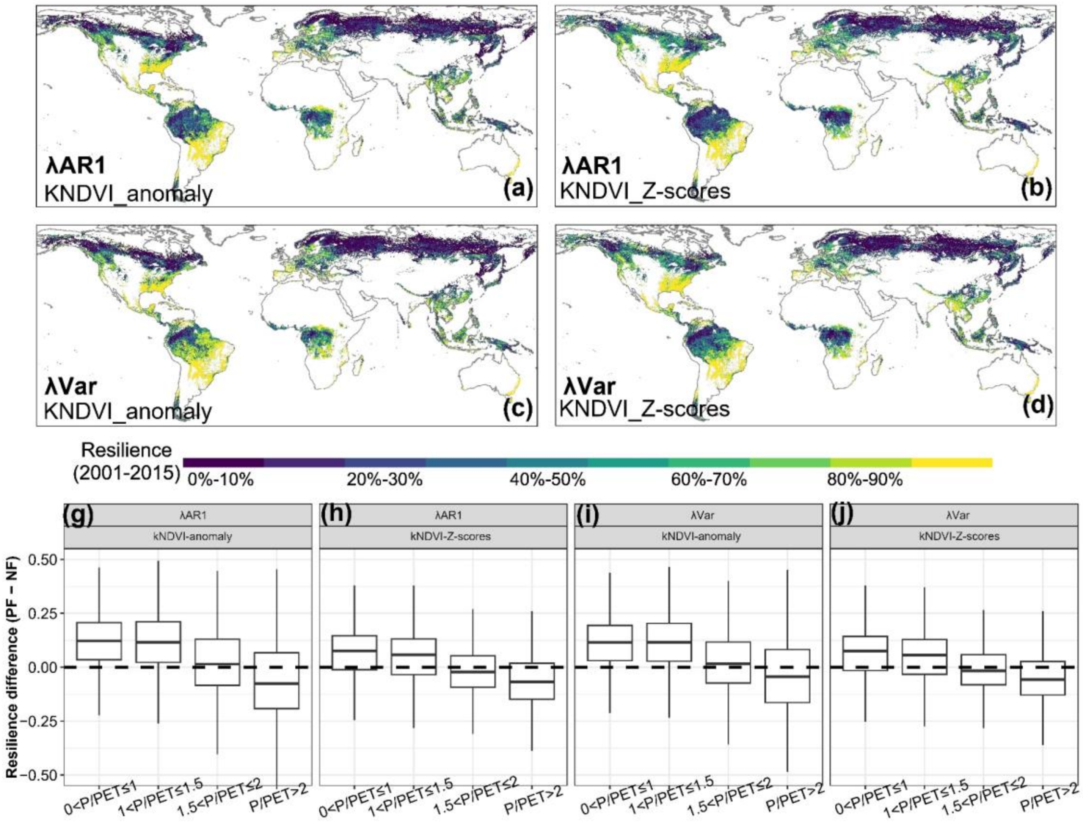


Figure R1. Sensitivity of resilience estimation to vegetation index standardization methods. Panels (a–d) show the global spatial distribution of resilience derived from λAR1 and λVar based on two kNDVI preprocessing approaches: (a, c) kNDVI anomalies (detrended deviations from the mean) and (b, d) kNDVI Z-scores (obtained by subtracting the long-term monthly mean and dividing by the corresponding monthly standard deviation). Colors indicate resilience percentiles during 2001–2015, where higher values represent lower resilience. Panels (g–j) summarize resilience differences (PF – NF) estimated using the two methods across precipitation-to-potential evapotranspiration (P/PET) zones for λAR1 (g, h) and λVar (i, j). The y-axis represents the resilience difference (PF – NF), calculated as the mean λAR1 or λVar in PF minus that in NF. Boxplots display the interquartile range, with the black dashed line indicating zero difference.

7. L99-102: what reviewer #1 originally asked was why you overlaid forest cover from 2000 with management from 2015. Is there no available forest management dataset for 2000? Reviewer #1 seemed to think that there is.

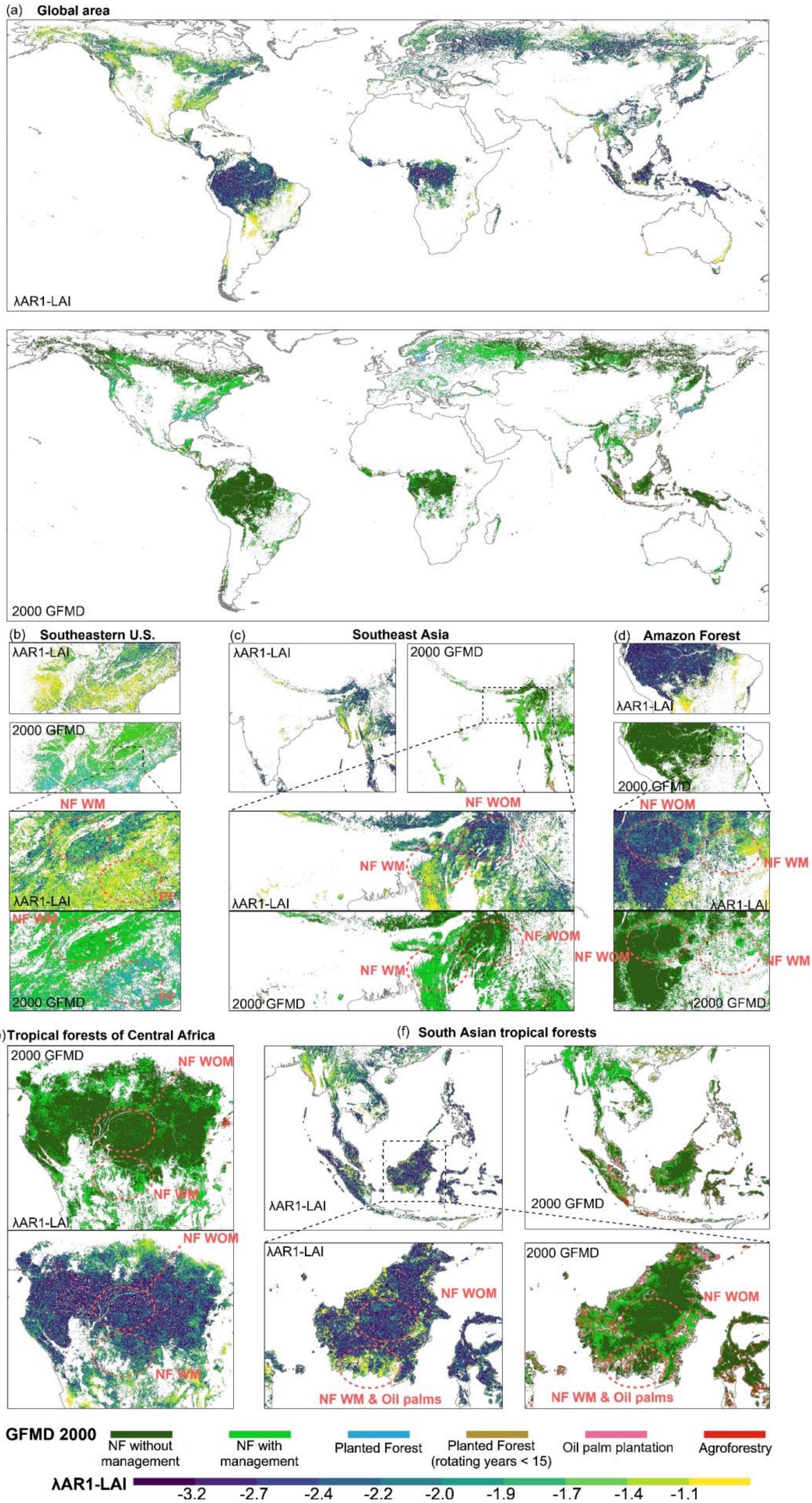
R: The global forest management dataset (GFMD) provided by Levi et al. contains data only for the year 2015, and no equivalent dataset is currently available for 2000. Thus, we overlapped the forest map in 2000 (i.e., 2000 global Hanssen forest map) and 2015 GFMD to generate the 2000 GFMD.

Moreover, many managed forests in 2015 GFMD, particularly planted forests with rotation cycles over 15 years (long-rotation planted forests), exhibit long-term and stable management practices. Therefore, we considered the 2015 forest management map to be a reasonable representation of persistent management patterns.

8. Maps in Fig. 1 and others: while maps provide a nice visualization and explicitly showcase spatial patterns, they can only be used for qualitative inspections of the data. So, I have a hard time agreeing with part of your interpretation. E.g.: "Results show that areas with human activity, as reflected in the forest management map, generally have lower resilience than nearby natural forests without management (NF WOM) (Fig. 1). This pattern is particularly evident in the Amazon and Central Africa, where the λAR1 estimated from LAI (λAR1-LAI) in natural forests with management (NF WM) is substantially higher than that in adjacent NF WOM (Fig. 1j, k and Supplementary Fig. S1).": the maps of South America in Fig. 1j and k can only be loosely compared and the only apparent differences where higher λAR1-LAI corresponds to managed NF is at the south-eastern edges of the Amazon forest, where e.g. edge effects might play a pivotal role. As for Central Africa, the maps in Fig. 1a and e are currently too low resolution to see any pattern, and the box in Fig. S1a-b cut out part of the central African forest.

R: We agree with the reviewer that Fig. 1 is intended for qualitative visualization rather than quantitative comparison. Our intention was to illustrate how forest management influences resilience within nearby regions, thereby providing a visual context for the subsequent quantitative analyses (i.e., moving-window results). The qualitative patterns in Fig. 1 suggest that forest management tends to weaken resilience locally, which motivated our paired sliding-window approach for quantifying the magnitude of such resilience differences.

Action Taken. Because online submission system may have caused a loss of image resolution. To address this, we have uploaded a high-resolution PDF version of Fig. 1/Fig. S1 to ensure that the spatial details are clearly visible. Additionally, we have provided regional zoom-in maps and global maps in Supplementary Fig. S1 to facilitate visual inspection. We used flesh circles and words to compare the differences in resilience under different forest management conditions (Fig. S1). Following the reviewer's suggestion, we have also revised Supplementary Fig. S1a–b to include the complete extent of the Central African tropical forests for a more comprehensive view. Based on these figures, we found that managed forests generally exhibit lower resilience than nearby unmanaged forests.



(a) Global area
λAR1-LAI
2000 GFMD
(b) Southeastern U.S.
λAR1-LAI
2000 GFMD
NF WM
PF
(c) Southeast Asia
λAR1-LAI
2000 GFMD
NF WOM
NF WM
(d) Amazon Forest
λAR1-LAI
2000 GFMD
NF WOM
NF WM
(e) Tropical forests of Central Africa
2000 GFMD
NF WOM
NF WM
λAR1-LAI
(f) South Asian tropical forests
λAR1-LAI
2000 GFMD
NF WOM
NF WM & Oil palms
GFMD 2000
NF without management
NF with management
Planted Forest
Planted Forest (rotating years < 15)
Oil palm plantation
Agroforestry
λAR1-LAI
-3.2
-2.7
-2.4
-2.2
-2.0
-1.9
-1.7
-1.4
-1.1

Supplementary Figure S1. Spatial distribution of forest resilience (λAR1–LAI) and management types from the Global Forest Management Dataset (GFMD 2000). Higher λAR1–LAI values indicate lower resilience. (a) Global distribution of forest resilience and management types. (b–f) Regional examples highlight contrasting resilience patterns among different forest management types. (b) In the Southeastern United States, planted forests (PF) exhibit lower resilience (higher λAR1–LAI) compared with nearby managed natural forests (NF WM). (c) In Southeast Asia, managed natural forests (NF WM) show lower resilience than unmanaged natural forests (NF WOM). (d) In the Amazon, reduced resilience is observed along forest edges where NF WM exhibit lower resilience than adjacent NF WOM. (e) In Central Africa, NF WM also show lower resilience than NF WOM. (f) In South and Southeast Asia, oil palm plantations and managed natural forests (NF WM & oil palms) exhibit substantially lower resilience than unmanaged natural forests. Flesh circles indicate representative regions used to illustrate contrasting resilience patterns. Color bars represent (top) λAR1–LAI values and (bottom) forest management categories from GFMD 2000, including NF without management (NF WOM), NF with management (NF WM), planted forest (PF), planted forest with rotation years <15, oil palm plantation, and agroforestry.

9. Fig. 2b, d, f: the figure information and label do not specify what is being confronted here, add specification at least to Fig. 2b label ("Latitudinal gradient of resilience difference between PF and NF […]").

R:**Action Taken.** We have revised the figure caption and labels to clearly state the comparison being presented. The updated caption now specifies that Fig. 2b, d, and f illustrate the *latitudinal gradient of resilience difference between PF and NF*. Specifically, the beginning of the relevant sentence in the caption now reads as follows:

"(b) Latitudinal gradient of resilience difference between PF and NF expressed in terms of λAR1–LAI for a set of five vegetation indices: GPP_MODIS (blue), GPP_PML (green), LAI (orange), kNDVI (blue-black), and FPAR (brown)." (Lines 177~179)

10. Fig. 3: the description of the axes and the title strip (with/without grey background) should be consistent (i.e., name the indicator consistently either in the axis title or in the facet title); d-e: add x-axis title to avoid confusion.

R: **Action Taken.** According to the reviewer's advice, we have revised Fig. 3 to ensure consistent labeling across all panels. The indicator names (λAR1, λVar, AR1) are now uniformly presented in the facet titles rather than mixed across axis and facet titles. In addition, we have added x-axis titles to panels (d) and (e) to clearly indicate the variable "P/PET" and avoid confusion.

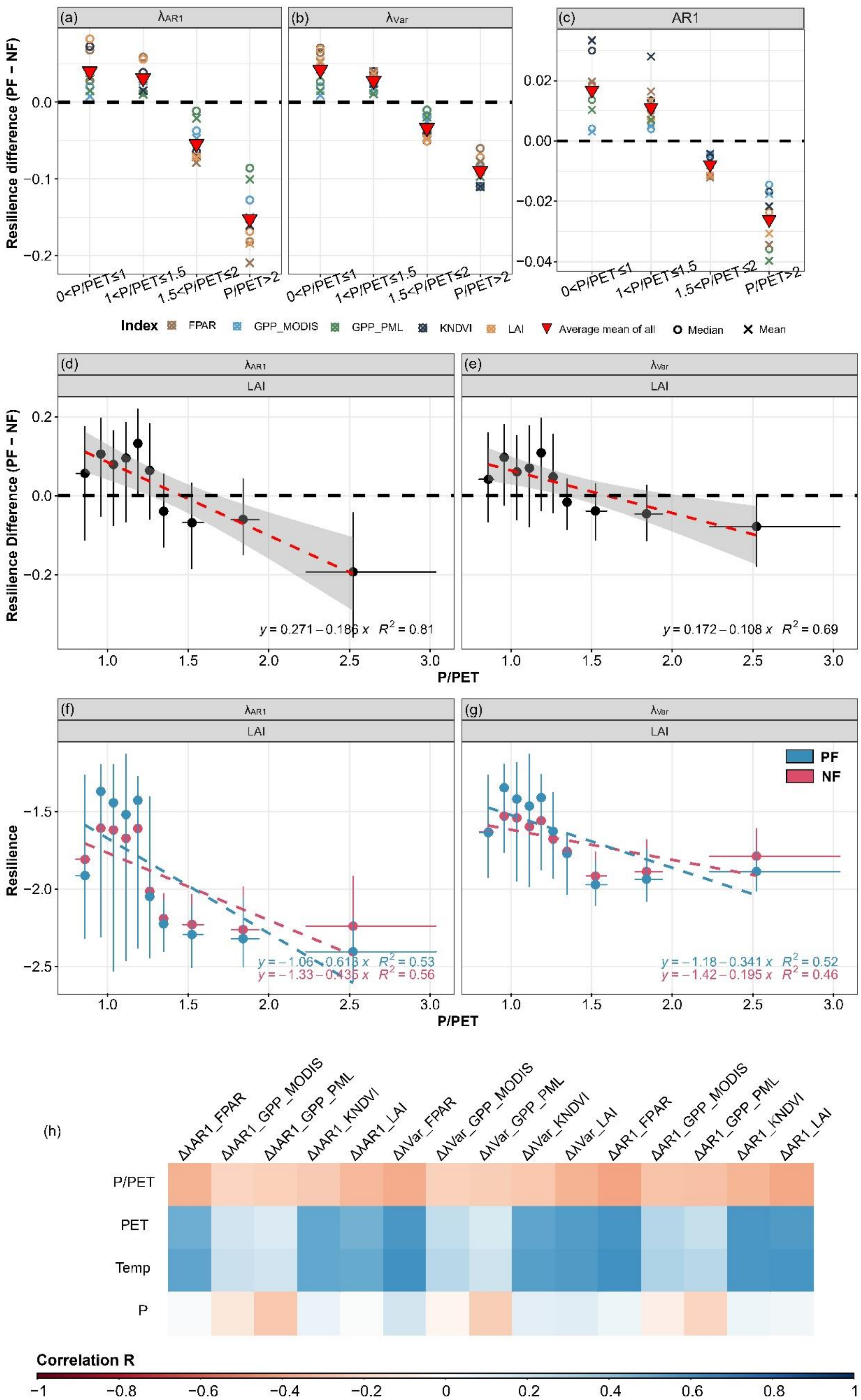


Fig. 3 Resilience difference (PF vs. NF) across a P/PET gradient. (a-e) Resilience difference

between PF and NF are shown across a P/PET gradient and explored using different CSD indicators and vegetation products as detailed in the legend. (f, g) Relationships between resilience and P/PET for individual PF and NF types, separately calculated from λAR1 and λVar. (h) Correlation matrix showing the relationship between resilience difference (PF vs. NF) and local biophysical factors, including the ratio of precipitation to potential evapotranspiration (P/PET), potential evapotranspiration (PET), precipitation (P), and temperature (Temp). Five vegetation indices and three CSD indicators are employed, resulting in 15 vegetation-resilience combinations. For example, "ΔAR1_FPAR" is the AR1 difference between PF and NF retrieved by FPAR. (a–c) A cross (×) marks the mean of each vegetation index, a circle (○) denotes the median, and a red inverted triangle (▼) indicates the average of the mean and median values. (d-g) Each point corresponds to the mean value within a P/PET decile at the global scale, with error bars indicating the 1st and 3rd quartiles.

Response to Reviewer #4

Reviewer #4 (Remarks to the Author):

The authors have provided a comprehensive explanation of the limitations of the study, in terms of variable input, analysis and future directions. They also have responded satisfactorily to my comments, especially presenting the magnitude of difference in the resilience between forest management types. Therefore, I approve of their revision but encourage the authors to consider highlighting the limitations of the study on the remote sensing and disturbance regime effect sides.

R: We sincerely thank the reviewer for the positive evaluation and constructive suggestions. Following your advice, we have highlighted the limitations related to remote sensing and disturbance regime effects in the Discussion section of the revised manuscript.

The limitations on the remote sensing side in influencing the data analysis and results are profound, but can be overlooked due to the global scale nature of the study. The spectral signatures of vegetation types and management conditions undergo constant change in strength and amount, following various natural and human disturbances in forest ecosystems, whether continuous, abrupt or subtle disturbances. So this in itself creates challenges for remote sensing data processing and analysis, in addition to image distortion challenges. So, I think the results offer a foundation for future studies using more fine resolution data such as from the global ecosystem dynamics investigation (GEDI). I encourage the authors to add these remote sensing related limitations in the manuscript to inform the readers of their awareness and transparency.

R: **Action Taken.** According to your suggestion, we have added the relevant content in the Discussion section as follows:

"Nevertheless, while this study integrates satellite observations to reveal management-related differences in forest resilience at the global scale, uncertainties may still arise from spectral variability, resolution constraints, and residual image distortions. Future integration of high-resolution structural information, such as that provided by the Global Ecosystem Dynamics Investigation (GEDI)[58], could further improve the reliability and precision of resilience assessments."

(Lines 508~513)

The limitations on the effect of specific disturbance type on the resilience of forest management types is also glaring, but dwelling on this in this study is unfeasible (it is a topic on its own), as such should be highlighted in the discussion and added as a recommendation for future studies. Different disturbance regimes affect different forest ecosystems differently in spatial and temporal scale.

R: **Action Taken.** According to your suggestion, we have added the relevant content in the Discussion section as follows:

"In addition, this study primarily focuses on long-term ecological resilience and does not explicitly examine specific disturbance events (i.e., engineering resilience)[36]. Future studies could therefore explore how specific disturbance events, such as droughts and fires, interact with forest management practices to influence recovery processes." (Lines 513~517)

References:

1. Holling, C. S. Surprise for Science, Resilience for Ecosystems, and Incentives for People. *Ecol. Appl.* **6**, 733–735 (1996).

2. Bathiany, S. *et al.* Ecosystem Resilience Monitoring and Early Warning Using Earth Observation Data: Challenges and Outlook. *Surv. Geophys.* **46**, 265–301 (2025).

3. Van Meerbeek, K., Jucker, T. & Svenning, J.-C. Unifying the concepts of stability and resilience in ecology. *J. Ecol.* **109**, 3114–3132 (2021).

4. Grimm, V. & Wissel, C. Babel, or the ecological stability discussions: an inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia* **109**, 323–334 (1997).

5. Studio, U. of E. D. Appendix 1: Glossary. *Global Tipping Points* https://report-2023.global-tipping-points.org.

6. Flores, B. M. *et al.* Critical transitions in the Amazon forest system. *Nature* **626**, 555–564 (2024).

7. Lovejoy, T. E. & Nobre, C. Amazon tipping point: Last chance for action. *Sci. Adv.* **5**, eaba2949 (2019).

8. Restrepo-Coupe, N. *et al.* Asymmetric response of Amazon forest water and energy fluxes to wet and dry hydrological extremes reveals onset of a local drought-induced tipping point. *Glob. Change Biol.* **29**, 6077–6092 (2023).

9. Brice, M.-H. *et al.* Moderate disturbances accelerate forest transition dynamics under climate change in the temperate–boreal ecotone of eastern North America. *Glob. Change Biol.* **26**,

4418–4435 (2020).

10. Zhang, Y. *et al.* Warming and disturbances affect Arctic-boreal vegetation resilience across northwestern North America. *Nat. Ecol. Evol.* **8**, 2265–2276 (2024).

11. Lunde, L. F. *et al.* Towards repeated clear-cutting of boreal forests – a tipping point for biodiversity? *Biol. Rev.* **100**, 1181–1205 (2025).

12. Sáenz-Romero, C. Arriving at a tipping point for worldwide forest decline due to accelerating climatic change. *For. Chron.* **100**, 5–7 (2024).

13. Hajdu, L. H., Meesters, A. G. C. A., Dolman, A. J. & Friend, A. D. Deforestation Could Push Amazonia Close to a Tipping Point Under Future Climate Change. *Geophys. Res. Lett.* **52**, e2024GL108304 (2025).

14. Terpstra, S. *et al.* Assessment of Abrupt Shifts in CMIP6 Models Using Edge Detection. *AGU Adv.* **6**, e2025AV001698 (2025).

15. Smith, T., Traxl, D. & Boers, N. Empirical evidence for recent global shifts in vegetation resilience. *Nat. Clim. Change* **12**, 477–484 (2022).

16. Boulton, C. A., Lenton, T. M. & Boers, N. Pronounced loss of Amazon rainforest resilience since the early 2000s. *Nat. Clim. Change* **12**, 271–278 (2022).

17. ENSO amplifies global vegetation resilience variability in a changing climate. https://www.researchsquare.com (2025) doi:10.21203/rs.3.rs-6534037/v1.

18. Wang, H. *et al.* Anthropogenic disturbance exacerbates resilience loss in the Amazon rainforests. *Glob. Change Biol.* **30**, e17006 (2024).

Reviewer #2 (Remarks to the Author):

The authors have satisfactorily responded to all my main comments. Before acceptance, I would only suggest the following minor changes:

R: We would greatly thank Reviewer #2 for the careful review and the positive assessment of our responses. The suggested minor changes have been fully addressed in the revised manuscript.

Supplementary L318-319: "PF generally possess lower structural complexity and shallower, simplified root systems", please add source(s) for this statement.

R: We have modified the related texts, now reading as: "Although PF and NF currently exhibit similar states in photosynthesis, PF generally possess lower structural complexity and shallower, simplified root systems[49]."

Supplementary Figures S22, S28...: you added the coefficient of variation, but labelled this as R instead of R2. If I am not mistaken, this should be R^2 (R-squared). Please check this and double check that it is kept consistent across all figures.

Similarly, the p-values are missing from e.g., Figures S28-S29.

R: We have checked and modified, please see the modified figures.

ROUND 3 REVIEWER 2 ATTACHMENT:

A lot of effort has gone in the rebuttal letter and in improving the robustness of the analyses and the ma in text. However, I still find that some issues were not fully answered and some clarifications still need to be included. Below my main concerns and some minor comments.

1. The definition of “**resilience**” seems to be a key point in your response, however, I found the definitions given to be unclear. Out of the sources you mentioned at the start of your response, as far as I can tell, four do not define resilience (Dakos et al., 2008; Lenton et al., 2012; Scheffer et al., 2009, 2012), three define resilience as variations of “return to the reference state after a disturbance” (Arnoldi et al., 2016; Grimm & Wissel, 1997; Verbesselt et al., 2016), i.e. what you define as engineering resilience, and only the remaining two references differentiate between engineering *vs.* ecological resilience, defining ecological resilience as the “magnitude of disturbance that can be absorbed before the system changes [to a different state]” (Bathiany et al., 2025; Holling, 1996). The subsequent sources you cite later in your response also only give mixed definitions that include both the ability to resist disturbances as well as the ability to recover from disturbances (Buschke & Brownlie, 2020; Forzieri et al., 2022; Smith et al., 2022). These two concepts (resist *vs* recover) should be differentiated more clearly. The way you describe ecological resilience as “concerning the system’s continuous response to a sequence of disturbances, both strong and weak” does not seem to reflect the cited references. Later in your response, you also mention that the indicators you are using “are also closely related to the recovery rate after disturbance”, making the whole concept even more unclear: if you are measuring the recovery rate after a disturbance, you seem to relate the study more to the engineering resilience, contrary to what you state.

   Aside from the somewhat confusing rebuttal letter, in the main text you define resilience as: “the capacity of ecosystems to withstand and recover from external perturbations while preserving their equilibrium state” (L72-74). This seems to be a mix between engineering (“recover”) and ecological (“withstand”) resilience, while the cited source differentiates between “the rate of recovery upon perturbation” (engineering resilience) and “the size of alternative attraction basins” (ecological resilience) (Scheffer et al., 2015). Finally, your definitions of resilience in the supplementary materials (“the ability of a system to persist and maintain its original state in the face of a continuous sequence of both strong and weak disturbance”) does not reflect the cited source, which only describes resilience as “the return rate from perturbations” (Boulton et al., 2022). At no point the cited references mention the concept of “continuous disturbance” or differentiate between “strong and weak disturbance”.

   Since the terminology around these concepts can be very confusing and has been used with different meanings in different contexts, it is important to be very clear about your definitions, keep them consistent across the whole text, and cite the correct sources. If I understand correctly, what you describe as “ecological resilience” could be termed “resistance” or “tolerance” as defined by Grimm & Wissel (1997), which is also in line with a more recent overview of the terminology by Van Meerbeek et al. (2021) (also *cf.* “latitude” as the “distance to threshold on the system variable axis”). These terms (e.g., resistance) might more intuitively convey what you are describing. You also use the term “recovery” almost interchangeably with “resilience”, but these two terms should be differentiated more clearly.

   Whatever terminology you choose to use be sure to define it and cite it correctly to avoid confusion, and keep it consistent throughout the text.

2. Definition of **tipping points**: you explain very in detail what tipping points are from a theoretical perspective. The addition of supplementary note 8 will definitely help clarify these notions to the readers. However, reviewer #1 asked specifically to define tipping points for your PF and NF. You only provide a general example where a forest turns into shrubland/grassland following severe a repeated disturbance. This does not differentiate between NF and PF and, importantly, it does not seem to align with the constraints set in your study (only forest areas). Within your study framework, are you seeing any such state shifts or tipping point? Is your filtering methodology intrinsically excluding any tipping point altogether?
3. **Temporal resolution** of indices: it is good that you included sensitivity tests. However, the provided figures do not give enough detail to fully evaluate these tests. E.g., Fig. S22 does not provide explained variance and goodness of fit metrics, and the colorbar is not described; Fig. S25 only show two regions (where I would argue that some differences can already be seen): what about other areas? Why show only these two?

   Moreover, you tested if the early warning indicators can detect changes in forest "resilience" prior to disturbance events (Fig. R4). Here, it is not clear 1) how you identified the two disturbance events; 2) how you tested for "significant" increases prior to the disturbance events; and 3) how you can state that "in both periods, the early warning indicators […] showed significant increases prior to the disturbances", while there is no calculated value for the indicators prior to 2003 (and thus prior to the first disturbance event shown in Fig. R4). On this note, in the main text you have to be careful about this distinction. E.g. at line 112-113 you state that "[CSD] indicators are computed over the entire observational period", after defining the observational period as 2001–2015. From my understanding, this is not true, since your 4-year moving window approach cuts the start and end of the timeseries. Please clarify.

   Finally, Fig. R5 is supposed to show increases in the indicators prior to a major disturbance event. Yet, the indicators show an increase throughout the whole period in which they were calculated, so how can this increase be related only to the disturbance event and rather than a longer-term trend?
4. **Spatial resolution**: reviewer #1 seemingly misunderstood this value as, clearly, 5 km ≠ 1 km, and your sensitivity analysis on the spatial resolution lends further robustness to your choice. Still, one nitpick in line with previous comments is that your provided evidence in Figure S29 is only based on qualitative visual analysis of the maps over African forests. Can you be sure that these two maps and the underlying data are statistically indistinguishable? And why are you only showing Africa? A suspicious reader might think that you are only showing what you want to show.
5. **Fig. S21** and others do not provide a sufficient explanation about what the y axis is showing and what statistical test was used to test the differences between management types. The reader has to refer to Figure 2, which makes the reading experience unnecessarily convoluted. Please add all relevant information in all figures' descriptions.
6. **Standardization**: in supplementary note 4 you mention that "Standardization of vegetation index data could influence resilience metrics. It is more likely to affect λVar, because standardization sets the standard deviation of the data to one.". That being said, you only tested the effect of standardization on λAR1. Here the appropriate sensitivity test would be to test it on λVar instead.

7. **L99-102**: what reviewer #1 originally asked was why you overlaid forest cover from 2000 with management from 2015. Is there no available forest management dataset for 2000? Reviewer #1 seemed to think that there is.
8. **Maps in Fig. 1 and others**: while maps provide a nice visualization and explicitly showcase spatial patterns, they can only be used for qualitative inspections of the data. So, I have a hard time agreeing with part of your interpretation. E.g.: "Results show that areas with human activity, as reflected in the forest management map, generally have lower resilience than nearby natural forests without management (NF WOM) (Fig. 1). This pattern is particularly evident in the Amazon and Central Africa, where the λAR1 estimated from LAI (λAR1-LAI) in natural forests with management (NF WM) is substantially higher than that in adjacent NF WOM (Fig. 1j, k and Supplementary Fig. S1).": the maps of South America in Fig. 1j and k can only be loosely compared and the only apparent differences where higher λAR1-LAI corresponds to managed NF is at the south-eastern edges of the Amazon forest, where e.g. edge effects might play a pivotal role. As for Central Africa, the maps in Fig. 1a and e are currently too low resolution to see any pattern, and the box in Fig. S1a-b cut out part of the central African forest.
9. **Fig. 2b, d, f:** the figure information and label do not specify what is being confronted here, add specification at least to Fig. 2b label ("Latitudinal gradient of resilience difference *between PF and NF* […]").
10. **Fig. 3**: the description of the axes and the title strip (with/without grey background) should be consistent (i.e., name the indicator consistently either in the axis title or in the facet title); d-e: add x-axis title to avoid confusion.

References:

Arnoldi, J.-F., Loreau, M., & Haegeman, B. (2016). Resilience, reactivity and variability: A mathematical comparison of ecological stability measures. *Journal of Theoretical Biology*, *389*, 47–59. https://doi.org/10.1016/j.jtbi.2015.10.012

Bathiany, S., Bastiaansen, R., Bastos, A., Blaschke, L., Lever, J., Loriani, S., De Keersmaecker, W., Dorigo, W., Milenković, M., Senf, C., Smith, T., Verbesselt, J., & Boers, N. (2025). Ecosystem Resilience Monitoring and Early Warning Using Earth Observation Data: Challenges and Outlook. *Surveys in Geophysics*, *46*(2), 265–301. https://doi.org/10.1007/s10712-024-09833-z

Boulton, C. A., Lenton, T. M., & Boers, N. (2022). Pronounced loss of Amazon rainforest resilience since the early 2000s. *Nature Climate Change*, *12*(3), 271–278. https://doi.org/10.1038/s41558-022-01287-8

Buschke, F., & Brownlie, S. (2020). Reduced ecological resilience jeopardizes zero loss of biodiversity using the mitigation hierarchy. *Nature Ecology & Evolution*, *4*(6), 815–819. https://doi.org/10.1038/s41559-020-1177-7

Dakos, V., Scheffer, M., van Nes, E. H., Brovkin, V., Petoukhov, V., & Held, H. (2008). Slowing down as an early warning signal for abrupt climate change. *Proceedings of the National Academy of Sciences*, *105*(38), 14308–14312. https://doi.org/10.1073/pnas.0802430105

Forzieri, G., Dakos, V., McDowell, N. G., Ramdane, A., & Cescatti, A. (2022). Emerging signals of declining forest resilience under climate change. *Nature*, *608*(7923), 534–539. https://doi.org/10.1038/s41586-022-04959-9

Grimm, V., & Wissel, C. (1997). Babel, or the ecological stability discussions: An inventory and analysis of terminology and a guide for avoiding confusion. *Oecologia*, *109*(3), 323–334. https://doi.org/10.1007/s004420050090

Holling, C. S. (1996). Engineering Resilience versus Ecological Resilience. In *Engineering Within Ecological Constraints*. National Academies Press. https://doi.org/10.17226/4919

Lenton, T. M., Livina, V. N., Dakos, V., van Nes, E. H., & Scheffer, M. (2012). Early warning of climate tipping points from critical slowing down: Comparing methods to improve robustness. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *370*(1962), 1185–1204. https://doi.org/10.1098/rsta.2011.0304

Scheffer, M., Bascompte, J., Brock, W. A., Brovkin, V., Carpenter, S. R., Dakos, V., Held, H., van Nes, E. H., Rietkerk, M., & Sugihara, G. (2009). Early-warning signals for critical transitions. *Nature*, *461*(7260), 53–59. https://doi.org/10.1038/nature08227

Scheffer, M., Carpenter, S. R., Dakos, V., & Nes, E. H. van. (2015). Generic Indicators of Ecological Resilience: Inferring the Chance of a Critical Transition. *Annual Review of Ecology, Evolution, and Systematics*, *46*(Volume 46, 2015), 145–167. https://doi.org/10.1146/annurev-ecolsys-112414-054242

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., van de Koppel, J., van de Leemput, I. A., Levin, S. A., van Nes, E. H., Pascual, M., & Vandermeer, J. (2012). Anticipating Critical Transitions. *Science*, *338*(6105), 344–348. https://doi.org/10.1126/science.1225244

Smith, T., Traxl, D., & Boers, N. (2022). Empirical evidence for recent global shifts in vegetation resilience. *Nature Climate Change*, *12*(5), 477–484. https://doi.org/10.1038/s41558-022-01352-2

Van Meerbeek, K., Jucker, T., & Svenning, J.-C. (2021). Unifying the concepts of stability and resilience in ecology. *Journal of Ecology*, *109*(9), 3114–3132. https://doi.org/10.1111/1365-2745.13651

Verbesselt, J., Umlauf, N., Hirota, M., Holmgren, M., Van Nes, E. H., Herold, M., Zeileis, A., & Scheffer, M. (2016). Remotely sensed resilience of tropical forests. *Nature Climate Change*, *6*(11), 1028–1031. https://doi.org/10.1038/nclimate3108