Reviewer Comments (RC) *are presented in black italics.*

Author Responses (AR) are presented in blue.

Changes in Manuscript (CM) are presented in green.

Response to Reviewer 1:

General Comments

RC1.1: *The manuscript by Wang et al. entitled "A Global Dataset of Forest Disturbance Regimes Derived from Satellite Biomass Observations" describes a novel global dataset quantifying forest disturbance regimes. The data are derived from inverse modeling constrained with data on global biomass and productivity data derived from remote sensing. The dataset is highly relevant, as disturbances remain ill constrained in global carbon cycle assessments and poorly incorporated in earth system models, partly because long-term disturbance data remain missing at global scale.*

I am impressed with this work and would really like to commend the authors for the general creativity and novelty of their approach. This has the potential to making a strong and important contribution to the field. That being said, there are a number of problems with the current manuscript which need to be addressed in a revised version of the work. I will describe my main concerns here, and give detailed line-level suggestions below.

AR1.1: We sincerely thank the reviewer for the highly positive assessment and encouragement regarding to the creativity, novelty, and potential impact of our framework on the field of global carbon cycle assessments. We also appreciate the constructive feedback, which has been instrumental in improving the clarity and ecological robustness of our manuscript. Detailed point-by-point responses to the Reviewer's concerns are provided below.

RC1.2: *First, and probably most important, it is not fully clear to me from the current text how exactly the dataset was derived. The flow chart in Fig. 1 helps, but the text remains vague and lacks details in many parts. Just to pick one example: How the random forest models were fitted, which variables were used as explained and explanatory variables, how random forest parameters were set etc. is not mentioned. Some of the technical details, e.g. addressing the scale mismatch and projection issues, are described with a high level of detail, but some of the more conceptual (ecological) aspects are rather glossed over. The problem here is that it is hard for the reader to understand how the dataset was generated, which limits confidence (particularly in combination with the issues with the evaluation presented – see below) in the dataset. I strongly suggest to revise this part and more clearly describe the process of deriving the dataset in the text.*

AR1.2: We agree with the reviewer that the derivation process of the dataset needed further clarification. To address this, we have updated the manuscript to detail the workflow, from its ecological assumptions to the specific machine learning configurations. Specifically, we made the following updates:

1. Workflow Overview in Introduction: We outlined the three main steps of our framework's implementation,

CM Line 101: "We achieved this by inverting forest disturbance framework... The implementation of this framework proceeded in three main steps: (1) expanding our previously established synthetic training dataset... to comprehensively represent realistic disturbance regimes by increasing parameter ranges and shapes of disturbances; (2) identifying the optimal spatial aggregation scale to bridge the scale mismatch... and (3) …evaluating the out of domain - extrapolation - predictions via the Dissimilarity Index (DIK)..."

2. Ecological mechanism in Section 2: We added paragraphs to explain the ecological mechanism, grounded in the "gap dynamics" concept.

CM Line 168: " Our framework aims to derive long-term disturbance regimes from a single, static realization of satellite forest biomass. This approach is grounded on the concept of gap dynamics: at a sufficiently large landscape scale (e.g., 25 × 25 km in our case), a forest ecosystem is continuously shaped by the interplay between carbon loss (disturbance events and background mortality) and carbon gain (gross primary productivity). Consequently, the landscape becomes a mosaic of individual patches in various stages of successional recovery (Watt, 1947; Bray, 1956; Bormann and Likens, 1979; Shugart, 1984; Bonan, 2015). Therefore, the spatial heterogeneity across the landscape acts as a substitute for time, capturing the region's history of disturbance and regrowth. ~~…~~A key challenge in inferring disturbance parameters from static biomass maps is equifinality... To overcome this, our framework relies heavily on second-order spatial texture features... By broadening the parameterization ranges... we scaled the number of simulated regime parameter combinations from 0.85 million to over 8 million... The pre-trained model subsequently integrates these observational spatial features with realistic GPP products... enabling a direct inversion of the underlying Disturbance Regime Parameters."

3. Machine Learning Details in Section 2.3: We described the Random Forest model configurations, including the explanatory and target variables, hyperparameter settings, HPC usage, and the cross-validation framework.

CM Line 413: " The inversion of forest disturbance regimes was formulated as a supervised regression task using the Random Forest algorithm in Python (scikit-learn). As explanatory variables, the models utilized 17 spatial-statistical metrics extracted from biomass maps (including first-order statistics, Shannon entropy, and GLCM textures) alongside FLUXCOM-X Gross Primary Production (GPP). The target variables were the four DRPs. For this production run, the models were trained on the entire synthetic dataset using all 17 features aggregated to the optimal 100 m x 100 m scale (kernel size of 10). We employed a standard ensemble configuration comprising 100 decision trees grown without depth restrictions and requiring a minimum of two samples to split an internal node. A fixed random seed was applied to guarantee reproducibility, and the training leveraged parallel processing across multi-CPU nodes on a high-performance computing (HPC) cluster. Model generalizability was robustly evaluated using a 10-fold cross-validation framework..."

Finally, we updated Figure 1 to visually align with this pipeline.

RC1.3: *Second, notwithstanding the merits of the author consortium in the fields of remote sensing and earth systems modeling, they need to read up a fair bit on disturbance ecology (i.e., the field they aim to contribute with their work here). This is particularly obvious by how their target variables are defined, which is not at all in line with the definitions and terminology commonly used in the field. Just to pick an example: What is described as intensity here does not describe disturbance intensity but rather disturbance severity. Also, for some of the parameters the units are missing or unclear (see more details below). These issues could lead to considerable confusion in the community and could strongly diminish the value of the work presented here. I hence suggest a revision of the terminology used and a more thorough description of the parameters (and their units) to make sure that the potential of the dataset will unfold to its full extent.*

AR1.3 (also addressing AR1.15-1.17): We thank the reviewer for these critical ecological distinctions. We fully agree that "intensity" refers to the physical energy of a driver, whereas "severity" describes the resulting ecological impact. Because our parameter (β) is a mathematical scaling slope governing the size-dependent fraction of biomass lost, it strictly measures disturbance severity, not intensity. We also agree that "extent" and "frequency" are better described as "disturbance rate" and "gap-size distribution."

To ensure rigorous terminology, we have globally replaced these terms and explicitly provided their exact definitions and units in Introduction and Section 3.1.1:

CM Line 97: "…The primary objective of this study is to globally map four key forest Disturbance Regime Parameters (DRPs): disturbance rate (μ [% yr-1], the mean annual fractional area affected), gap-size distribution (α [-], the scaling governing the spatial clustering of disturbance patch sizes; Fisher et al., 2008), disturbance severity (β [-], the scaling slope governing the size-dependent fraction of biomass lost; Chambers et al., 2013), and background mortality (Kb [yr-1], the continuous non-episodic carbon turnover)."

CM Line 465: "…

- Disturbance Rate (m [% yr-1]): the mean annual fractional area affected, distinguishing between highly disturbed and stable regimes.
- Gap-size Distribution (a [-]): the scaling exponent governing the spatial clustering of disturbance patch sizes, distinguishing between regimes of many small events versus few large events.
- Disturbance Severity (b [-]): the scaling slope governing the size-dependent fraction of biomass lost, distinguishing between regimes of low-severity (e.g., partial canopy thinning) and high-severity (e.g., stand-replacing).
- Background Mortality (Kb [yr-1]): the baseline mortality rate from non-episodic processes of carbon turnover, such as natural decay and competition.

  …"

RC1.4: *A third issue I see is with using GPP data only for a single year (2010). This means that the idiosyncratic weather patterns of this particular year (e.g., drought in one place, heat wave in another, extensive rain in yet another region) will be baked into the data. This, in turn, does not correspond well* *with the biomass data used, which is the integral over many years or decades (and in some cases centuries). I suggest to use a multi-year average GPP estimate here to increase the robustness of the assessment.*

AR1.4: We agree with the reviewer that relying on a single year of GPP data could introduce biases from short-term weather anomalies. To address this concern and increase the robustness of our assessment, we have rerun the entire pipeline and the following analysis using a multi-year mean of GPP spanning the decade from 2001 to 2010. Furthermore, to verify the stability of the spatial patterns, we conducted a grid-to-grid comparison between the original 2010 data and the updated 2001-2010 mean, which is now included in the Supplementary Materials (S4.3). This comparison confirms that while the macro-scale spatial gradients of GPP are generally stable over time, utilizing the 10-year average explicitly mitigates transient interannual variability and short-term climatic noise as the reviewer suggested. We have updated the Methods section to reflect the use of the multi-year mean and added the comparison figure to the supplement.

Some more observations:

RC1.5: *I strongly suggest to somewhere early on state which disturbances you address here. From my understanding, it is all canopy mortality events, including both human and natural disturbances. While any disturbance that only affects the understory but not the overstory (e.g., low-severity fire) and hence has no strong signature in aboveground biomass is not considered. This is me guessing what I think it is, but readers should not need to guesstimate what the target variable of the assessment is. Please be specific here!*

AR1.5: We thank the reviewer for highlighting the necessity of an explicit definition. We agree that defining our target variable early in the text is crucial for clarity. To ensure this is clear from the onset of the paper, we have updated the beginning of the second paragraph in the Introduction. We now explicitly differentiate between a "disturbance" (the event) and the "disturbance regime" (the long-term pattern).

CM Line 66: “In the context of this study, a forest disturbance is explicitly defined as any event, whether natural or anthropogenic, that results in a measurable reduction of above ground biomass (AGB), thereby leaving a structural spatial signature on the forest landscape. Consequently, forest disturbance regimes describe the long-term spatial and temporal patterns of these biomass-reducing mortality events within a landscape, encompassing varied natural and anthropogenic processes that result in a significant loss of aboveground biomass.”

RC1.6: *I am somewhat missing a more reflected discussion of limitations. The spatial aggregation approach undertaken is mainly needed because the underlying model is lacking the representation of important spatial processes (contagion). This is indeed cleverly done here and makes sense for the product. But one could also argue that this approach (which reduces the fidelity of the dataset overall by* *the need to spatially average) would not be needed if a better/ more appropriate/ more refined underlying model would have been used. Also, the evaluation basically tests against data that were used for generating the products, so is not an evaluation against independent data. While I understand that this is difficult to do, some tests against local data on disturbance regimes (from dendroecological sources, long-term inventory data, etc.) would have been desirable. Also, while the approach taken here is generally clever and creative, it remains unclear to me how the effects of the different aspects of disturbance can be teased apart with high certainty, given that at pixel level, disturbance rate, size, and severity will result in a similar effect on biomass (reduction). Early on in the text the authors claim that their model can do this well, but why it can and how well it actually can remains unclear for me. I am not saying that the authors should fundamentally change these things (as this will probably not be possible and would substantially change their work), but a somewhat more reflected and nuanced discussion of the limitations of their work would be highly appreciated.*

AR1.6: We appreciate the reviewer's insightful comments regarding the limitations of our approach. We agree that a more nuanced and reflective limitation discussion is necessary. To address these concerns, we have expanded our manuscript across three specific dimensions to ensure transparency:

1. Equifinality: The reviewer points out that different disturbance parameters can result in similar mean biomass reductions at the pixel level. However, our framework does not rely on mean biomass alone. As demonstrated in our synthetic evaluations (Wang et al., 2024), different disturbance regimes (e.g., frequent small gaps vs. infrequent large clearings) produce mathematically distinct biomass patterns and distributions. By utilizing 17 high-dimensional spatial-statistical features (such as GLCM texture and variance), the machine learning model can effectively disentangle these effects. We have explicitly clarified this mechanism in the Methods section (see AR1.2 Ecological mechanism in Section 2).

2. Independent Evaluation: We fully agree that evaluating against independent data would further strengthen the dataset. To test the reliability of our product, we conducted a preliminary cross-biome validation across European regions using independent, Landsat-derived forest disturbance records (1985–2010). Although the two frameworks exhibit fundamental operational differences, such as mismatched temporal windows and the fact that the Landsat imagery primarily captures distinct stand-replacing events, the comparison indicates that the retrieved parameters fall within a highly comparable range to those estimated from the event-based observations across major biomes. The detailed validation methodology and the corresponding results have been added to Supplement Section S4 (visualized in Figure S4.6). However, because a comprehensive, globally continuous independent validation remains an ongoing research effort, we maintain a transparent stance and explicitly discuss the current lack of a global independent validation as a key limitation in Section 3.3 and the Conclusion.

CM Line 617:” Nevertheless, a preliminary cross-biome comparison over European forests using independent Landsat-derived records shows that our retrieved parameters fall within a highly comparable numerical range to event-based observations (Figure S4.4).”

3. Model Simplicity and Spatial Contagion: We fully agree with the Reviewer's assessment. If our framework employed a more complex mechanistic model that explicitly resolved the physical propagation of disturbances, it would inherently generate more realistic spatial autocorrelation. We have now acknowledge this limitation in the manuscript (Discussion 3.4.1), clarifying that the spatial aggregation step mitigates the simplicity of our current event generator (Method 2.2).

CM Line 317 Method 2.2:"The primary source of inconsistency was identified in the representation of spatial patterns: the simulation framework treats each grid cell as an independent unit, whereas real-world landscapes exhibit strong spatial autocorrelation arising from complex interaction among topography, soil, hydrology, community competition and the contagious nature of disturbances, such as fires spreading through landscapes. As such, spatial aggregation likely mitigates the simplicity of the approach used here for generating disturbance events, which may fall short in generating realistic spatial patterns by not resolving mechanistically disturbance spread."

CM Line 695:"3.4.1 Pattern-based Inference versus Process-based Simulation

It is important to distinguish the conceptual scope of our framework from dynamic process-based models. In disturbance ecology, dynamic process-based models are designed to mechanistically simulate the physical spread of events over time, such as the step-by-step propagation of fire or insect outbreaks (e.g., Rammer and Seidl, 2022). In contrast, our approach operates within a pattern-based statistical paradigm. Consequently, our framework focuses on evaluating the emergent spatial attributes of disturbed landscapes, namely the size, form, and distribution of biomass loss, rather than simulating the mechanistic temporal spreading processes (i.e., spatial contagion) themselves.

Within this pattern-based paradigm, the stochastic disturbance event generator places independent geometric shapes onto the simulated landscape. Crucially, while these synthetic generative grids are inherently abstract without a predefined physical resolution, real-world EO biomass products possess strict, specific spatial resolutions. To bridge this fundamental scale mismatch and align the discrete simulated events with the continuous patterns of satellite data, spatial aggregation (Section 2.2) is utilized as a necessary scaling step to anchor the spatial statistics to the observational scale, rather than serving as a correction for model deficiency.

Furthermore, instead of explicitly simulating mechanistic contagion, our framework mathematically encapsulates the structural outcomes of complex spatio-temporal interactions, including linked disturbances and disturbance cascades, through a massive training library of over 8 million diverse regime combinations. By incorporating highly clustered gap-size distributions (α) and a wide diversity of complex, non-rectangular disturbance morphologies, the training dataset statistically approximates the diverse spatial footprints left by contagious processes. This pattern-based representation provides a highly efficient and robust means to infer generalized disturbance regime parameters from static landscape biomass patterns at a global scale."

More detailed line-level comments:

RC1.7: *L20: not only number of disturbances, but also size and severity are changing, and contribute equally to the issue*

AR1.7: We have revised the corresponding sentence in the Introduction:

CM Line 20: "Forests play a central role in the global carbon cycle by serving as critical carbon sinks for atmospheric CO_2. Yet, the stability and continued capacity of these sinks are increasingly threatened by the growing number, size, and severity of disturbances."

RC1.8: *L34: borderline disturbance regimes – meaning unclear*

AR1.8: We agree that the term "borderline" was ambiguous. To ensure terminological precision, we have replaced it with **CM Line 35**: "an empirical evaluation comparing paired forest landscapes with contrasting extremes in their predicted disturbance parameters supports the methodology and assumptions used to build the dataset."

RC1.9: *L54: change detection*

AR1.9: Corrected.

RC1.10: *L70: I am not familiar with the term "forward-modeling", which is used quite frequently throughout the text; please explain at first usage what you mean*

AR1.10: In our context, "forward-modeling" refers to the process of dynamically simulating the expected spatial biomass patterns (the observable outcomes) from a prescribed set of disturbance parameters (the causal mechanisms). We have now explicitly defined this term at its first usage in Section 2.

CM Line 176: "we utilized a forward-modeling approach (Wang et al., 2024), that generates the AGB patterns resulting from the modelled dynamics of forest carbon coupled to a stochastic disturbance-events generator. We use the inverse modeling approach, based on a Random Forest that determines from the emerging AGB patterns and distributions.."

RC1.11: *L129: how is the approach "adaptive"?*

AR1.11: We thank the reviewer for pointing out this ambiguity. By "adaptive," we meant that our machine learning framework learns how to inter the disturbance regime parameters directly from the emergent spatial patterns in the biomass data. To eliminate any confusion, we have removed the ambiguous term and revised the relevant sentence to clarifying this meaning.

CM Line 101: "We achieved this by inverting an adaptive forest disturbance framework, which learns how to infer the long-term DRPs directly from the emergent spatial patterns of satellite biomass data."

RC1.12: *L131: I suggest that all 17 features are also explicitly named in the main text, maybe move the table from the supplement to the main?*

AR1.12: We thank the reviewer for the suggestion. To keep the main text concise, we keep the full detailed table in the supplement, but also add a simple summary table (Table 1) in Section 2.1 to explicitly list and categorize all 16 biomass features.

CM Line 225: " From this masked AGB grid, following Wang et al. (2024), we calculated a comprehensive suite of biomass spatial statistics (Table 1), including first-order distribution metrics, an informative feature, and second-order texture metrics from a Gray-Level Co-occurrence Matrix (GLCM). Detailed explanations for each feature can be found in Supplementary Table S3."

Table 1. Summary of the 16 spatial biomass statistical features

Feature Type	Specific Feature Names
First-order Histogram Features	Mean, Median, Variance, Standard Deviation, Coefficient of Variation, Skewness, Kurtosis, 25th Percentile, 75th Percentile, Range, Trimean
Informative Feature	Shannon Entropy
Second-order Texture Features	GLCM Contrast, Correlation, Energy, Homogeneity

RC1.13: *L145: is it robust to use the GPP of a single year here, given that the biomass data are the aggregate over a much longer time horizon? The year selected could have been a particularly dry year in some places, while not in others, which will affect the GPP, and hence the particular weather patterns of that year will be represented in the GPP layer (while the biomass pattern is the integral over decades to centuries).*

AR1.13: We fully agree with the reviewer. We have addressed this exact concern by replacing the single-year GPP with a 10-year multi-year mean (2001-2010). Please refer to our detailed response and the new supplementary figure in AR1.4.

RC1.14: *L160: and the contagious nature of disturbances, such as fires spreading through landscapes*

AR1.14: We agree, the exact phrasing was added.

CM Line 318: "…whereas real-world landscapes exhibit strong spatial autocorrelation arising from complex interaction among topography, soil, hydrology, community competition and the contagious nature of disturbances, such as fires spreading through landscapes. As such, spatial aggregation likely mitigates the simplicity of the approach used here for generating disturbance events, which may fall short in generating realistic spatial patterns by not resolving mechanistically disturbance spread."

RC1.15: *L233: is the base for calculating extent forest area or total cell area? This should be specified. Also, I suggest to rename this to "disturbance rate", as in disturbance ecology, extent is commonly an* *estimate of area affected (e.g., in ha or km²), while the annualized relative area affected (which is what you report here) is referred to as rate*

AR1.15: Agree and noted in AR1.3.

RC1.16: *L234: frequency is not a pattern in space, but a pattern in time. Whether the events are small or large is best described by gap size distribution; also unclear: what is the unit here? Without this it is very hard to interpret.*

AR1.16: Agree and noted in AR1.3.

RC1.17: *L236: The term severity and intensity are used interchangeably here, yet they refer to distinctly different properties of the disturbance regime. Please don't use them interchangeably, as this will only confuse readers and decrease the value of your work for the community (see e.g. Turner 2010, Ecology 91, 2833-2849, Table 1 for definitions of key terms of disturbance ecology). What you report, based on your description, is severity, so please also call it that (and not intensity, which describes the energy of an event, e.g. the fire line energy, or the gust wind speed of a storm). And same comment on unit here: Should be % or rate [0,1], I assume, but is not specified explicitly – please clarify!*

AR1.17: Agree and noted in AR1.3.

RC1.18: *L238: What is the unit exactly here? I.e., what per year is reported? Biomass loss in g/m²? Percent biomass loss? Other? Please be explicit, as otherwise it is difficult for the community to understand (and use) your product!*

AR1.18: Agree and noted in AR1.3.

RC1.19: *L332: I agree when it comes to the global extent, but there are many regions of the world where these properties are well understood and where we know e.g., from dendroecological studies, long-term inventories etc. about the properties of forest disturbance regimes. So the statement is not entirely correct.*

AR1.19: We agree with the reviewer. Validating our results against regional data, such as dendroecological studies and forest inventories, is an important direction for our future work. We have revised the text and added this to our outlook in the Conclusion.

CM Line 615: "A direct, quantitative validation of our disturbance regimes is inherently challenging. While regional properties are documented through dendroecological studies and long-term inventories, a globally consistent dataset at this landscape scale is currently missing."

CM Line 868: "…and the current lack of extensive, independent global ground-truth validation. Addressing these gaps, particularly by integrating and comparing our dataset with regional observations, represents critical directions for future research. …"

RC1.20: *L334: Well, you used this data to produce the maps, so comparing the data you generated back to the input data you used to generate them is not a very strong test – it is not surprising at all, that you find patterns here! An independent test, even if only for some selected regions, would be much stronger.*

AR1.20: We agree that an independent test is the most intuitive and persuasive approach, and we have explicitly acknowledged this lack of independent validation as a limitation (see AR1.6). However, we must clarify that this section is designed as an empirical plausibility check. Its specific purpose is to verify whether real-world landscapes with obvious structural differences in biomass are successfully and accurately differentiated by our predicted parameters. We have now explicitly clarified this objective in Section 3.3.

CM Line 619: "Therefore, in this section we assessed the plausibility of our product by verifying whether real-world landscapes at the extreme tails of the global DRPs distributions (i.e., the lowest and highest predicted values) are successfully and accurately differentiated by our framework. To isolate the impact of each disturbance parameter (μ, α, β), we selected paired contrasting landscapes (Figure 7c) that shared highly similar background conditions (GPP, Kb) and identical non-target disturbance parameters, thereby allowing a controlled visual assessment of how each parameter uniquely influences biomass patterns (Figure 7)."

RC1.21: *L335: unclear, what is a scenario in your context? As far as I can see, you did not conduct any scenario analyses. Consequently, I don't understand what extreme high/low scenarios are, high/ low with regard to what?*

AR1.21: We agree that the term "scenario" was misleading, as we did not conduct any future or alternative scenario analyses. By "extreme high/low scenarios," we simply meant the extreme upper and lower tails of our globally predicted parameter distributions. To eliminate this confusion, we have removed the word "scenario" throughout this section and clarified the phrasing, as detailed in the revised text provided in AR1.20.

RC1.22: *L359:* shows *a more transition… meaning unclear*

AR1.22: We intended to describe a landscape with more gradual, continuous spatial changes in biomass, as opposed to the sharp, distinct boundaries left by high-severity disturbances. We have rewritten this by:

CM Line 646: "…as low- β shows a more gradual spatial transitions in biomass while high- β features distinct boundaries typical of severe disturbance footprints."

RC1.23: *L361: I think the term "scenario" is used wrongly here, see also my comment above*

AR1.23: The term "scenario" was misleading. We have replaced "scenarios" with "characteristics" in this sentence.

CM Line 664: “It underscores the necessity of using spatial-statistical features to capture more sophisticated disturbance characteristics, validating our model’s approach.”

Response to Reviewer 2:

General comments

RC2.1: *This paper attempts to produce representative values of 3 disturbance parameters and background mortality for forests within 25 km x 25 km tiles and 0.25° x 0.25° cells using forward simulations from a model that is massaged to give comparable values across 17 measures to those found using the 2010 GlobBiomass AGB map. It is very much a paper produced by a computer scientist in which the physical reality of disturbance is hardly considered. As a symptom of this, no definition of forest disturbance is given in the paper. Is it deforestation, degradation (in its many forms); does forest growth come into it? Similarly, no definition is given of the 3 disturbance parameters. For example, what is "extent" of disturbance in a possible mix of forest clearance and degradation? How is severity defined? And what is frequency (is this temporal or spatial frequency)? Bizarrely, the three parameters (μ, α, β) have different meanings in the main text and supplement. How is disturbance distinguished from natural variation in space?*

AR2.1: We sincerely thank the reviewer for this honest feedback. As also noted by Reviewer 1 (RC1.3 and RC1.5), certain terminology and imprecise descriptions of ecological mechanisms could be misleading to readers. To address this, we have thoroughly revised the relevant content throughout the entire manuscript:

1. Explicit Definitions Added: We have added clear definitions of forest disturbance and long-term disturbance regimes at the beginning of the Introduction to clarify the physical and ecological context of our study.
   CM Line 66: "In the context of this study, a forest disturbance is explicitly defined as any event, whether natural or anthropogenic, that results in a measurable reduction of above ground biomass (AGB), thereby leaving a structural spatial signature on the forest landscape. Consequently, forest disturbance regimes describe the long-term spatial and temporal patterns of these biomass-reducing mortality events within a landscape, encompassing varied natural and anthropogenic processes that result in a significant loss of aboveground biomass."

2. Harmonized Parameter Terminology: We have globally updated the parameter names across the main text, figures, and supplementary materials to "disturbance rate (μ)", "gap-size distribution (α)", and "disturbance severity (β)" to eliminate prior inconsistencies between sections.

   CM Line 96: "…four key forest Disturbance Regime Parameters (DRPs): disturbance rate (μ [% yr-1], the mean annual fractional area affected), gap-size distribution (α [-], the scaling governing the spatial clustering of disturbance patch sizes; Fisher et al., 2008), disturbance severity (β [-], the scaling slope governing the size-dependent fraction of biomass lost; Chambers et al., 2013), and background mortality (Kb [yr-1], the continuous non-episodic carbon turnover…"

3. To clarify how our framework isolates disturbance signatures from natural spatial variation, we have expanded Section 2 to explicitly detail the "shifting mosaic steady-state" concept. By combining high-dimensional spatial texture features with GPP data, the model resolves equifinality and distinguishes true disturbance footprints from baseline natural variations.

   CM Line 168: “Our framework aims to derive long-term disturbance regimes from a single, static realization of satellite forest biomass. This approach is grounded on the concept of gap dynamics: at a sufficiently large landscape scale (e.g., 25 × 25 km in our case), a forest ecosystem is continuously shaped by the interplay between carbon loss (disturbance events and background mortality) and carbon gain (gross primary productivity). Consequently, the landscape becomes a mosaic of individual patches in various stages of successional recovery (Watt, 1947; Bray, 1956; Bormann and Likens, 1979; Shugart, 1984; Bonan, 2015). Therefore, the spatial heterogeneity across the landscape acts as a substitute for the temporal changes, capturing the region's history of disturbance and regrowth.

   To quantitatively link these spatial biomass patterns to their underlying disturbance regimes, we utilized a forward-modeling approach (Wang et al., 2024), that generates the AGB patterns resulting from the modelled dynamics of forest carbon coupled to a stochastic disturbance-events generator. We use the inverse modeling approach, based on a Random Forest that determines from the emerging AGB patterns and distributions. A key challenge in inferring disturbance parameters from static biomass maps is equifinality, that is different combinations of disturbances can produce similar mean biomass outcomes. To overcome this, our framework relies on second-order spatial texture features rather than solely on first-order statistics. For this study, we significantly advanced our previous synthetic training library. By broadening the parameterization ranges and incorporating a diversity of non-rectangular disturbance shapes, we scaled the number of simulated regime parameter combinations from 0.85 million to over 8 million. By forward-simulating millions of forest landscapes under diverse disturbance regime parameter combinations, we generated a massive training dataset that uniquely links specific disturbance regimes to their resulting spatial biomass patterns (Supplementary S1).

   In application, rather than relying on landscape biomass averages, the framework disentangles this equifinality by extracting a suite of biomass spatial statistics (such as variance, skewness, and GLCM texture) from the observed satellite biomass map. These statistical features serve as direct proxies for structural complexity. For instance, high GLCM homogeneity corresponds to large, even-aged stands created by infrequent, large-scale, stand-replacing disturbances. Conversely, high spatial variance or GLCM contrast reflects a multi-aged, structurally complex forest driven by frequent, small-scale gap dynamics. The pre-trained model subsequently integrates these observational spatial features with realistic GPP products, which constrain the post-disturbance growth and recovery rates, enabling a direct inversion of the underlying DRPs.”

RC2.2: *However, the major problem with the paper is that it relies uncritically on the GlobBiomass map as being a faithful representation of disturbed forest regions without paying any attention to the limitations of the map, including biases, saturation effects, large dispersion at 25 m and possible spatial correlation induced by the data processing (which may have nothing to do with spatial correlation in AGB). There is no discussion of the qualities of the map and how they might affect its ability to truly represent the properties of disturbed forests. Hence the "disturbance" they are representing is entirely a description of pixel properties of the map, which may, inter alia, include false evidence of disturbance or fail to capture true disturbance.*

AR2.2: We acknowledge that these limitations exist within satellite-derived biomass products and can propagate into downstream parameter inversions. To address these concerns and transparently discuss the boundaries of our dataset, we have constructed a new dedicated section, Section 3.4 ("Methodological Scope and Limitations"), which incorporates extensive supplementary analyses to complement and illustrate:

1. Pattern-based Inference versus Process-based Simulation (Section 3.4.1): We clarify the conceptual boundaries of our framework, which operates within a pattern-based statistical paradigm rather than simulating temporal spreading processes mechanistically. We explain that spatial aggregation is utilized as a necessary scaling step to bridge the scale mismatch between abstract simulations and real-world observations.
2. Sensitivity to the Choice of Biomass Product (Section 3.4.2): We systematically evaluated the framework's robustness by repeating the inversion pipeline across two alternative ESA CCI Biomass products (2010, and a multi-year mean). The results demonstrate that the retrieved macro-scale geographic gradients remain highly consistent and stable regardless of the specific baseline input product used.

Specific comments

RC2.3: *The AGB map is for 2010 but the landcover is for 2019. How does this affect the results?*

AR2.3: We agree with the reviewer that a temporal mismatch existed in the previous version. To address this and ensure spatiotemporal consistency, we have updated the entire workflow. For the 2010 baseline above-ground biomass (AGB) map, we now consistently use the 2010 global land cover map (CCI Land Cover 2010). For multi-year datasets, such as the ESA CCI Biomass v7 mean, we adopted a dynamic land cover approach, aligning the land cover mask of each specific year with the corresponding biomass observations to accurately account for concurrent forest transitions.

CM Line 260: "The forest cover mask used in this study was derived from the European Space Agency Climate Change Initiative (ESA CCI) Land Cover dataset (1992–2015), which categorizes land surfaces based on the UN-FAO Land Cover Classification System (LCCS). The land cover layers were reprojected and resampled to match the spatial extent and resolution of each biomass tile using nearest-neighbor interpolation. We then identified pixels belonging to tree cover categories (specifically LCCS codes 50–

90, 160, and 170, which encompass evergreen needleleaf, evergreen broadleaf, deciduous needleleaf, deciduous broadleaf, mixed, and flooded forests) to generate a binary forest mask, while excluding missing data.

Crucially, this binary mask serves two fundamental purposes in our framework: first, to filter out entirely non-forest landscapes (retaining only landscape tiles with a forest cover fraction greater than 0); and second, to strictly isolate forest-covered pixels within those retained landscapes, ensuring that the extraction of spatial biomass statistics and the subsequent prediction of disturbance regimes are computed exclusively using actual forest biomass values.

To align with our modeling objectives, we implemented a specific temporal matching strategy for the masks to correspond with the timeline of each input biomass dataset. For the static, single-year baseline biomass analysis, we consistently applied the fixed 2010 land cover layer to match the 2010 observation snapshot. In contrast, for the 2005–2011 annual biomass time series, we dynamically aligned the land cover mask with the corresponding year of the biomass observations (e.g., pairing the 2005 biomass data with the 2005 land cover layer) to accurately account for concurrent forest transitions."

RC2.4: *The phrase "potential extrapolation" is used in several places without ever being properly explained. The idea seems to be that if the simulation disagrees with the properties of the map then it's because the training set isn't large enough, but there's no evidence to support this. Why can't it be because the trained data just makes mistakes?*

AR2.4: We thank the reviewer for highlighting the need to clarify the term "potential extrapolation." To directly address the concern of whether the model is simply "making mistakes," we clarify that our framework explicitly separates inherent model prediction errors from true statistical extrapolation. We have restructured the text to make this distinction clear:

1. Defining Extrapolation: In our context, extrapolation has a strict statistical definition: it occurs when the structural features of an observed satellite landscape (e.g., unusual textures resulting from human management or sensor artifacts) fall entirely outside the multi-dimensional feature space of our theoretical training set. To clarify this, we updated the Introduction to explain that **CM Line 105:** " identifying the optimal spatial aggregation scale to bridge the scale mismatch between simulations and EO product, explicitly reducing the risk of model extrapolation (i.e., applying the model to spatial feature value ranges not covered by the synthetic training data)";

2. Distinguishing "Mistakes" from Extrapolation: We provide two distinct uncertainty metrics to independently quantify these issues:

> Model Error: The internal prediction uncertainty, when the model makes errors within its known domain, is quantified by the standard deviation across the 10 cross-validation folds (std_all_folds).

Extrapolation: The applicability uncertainty is independently quantified using the Dissimilarity Index (DIK). **CM Line 428:** "DIK values below 1.0 suggest the landscape is well-represented, whereas values significantly greater than 1.0 serve as a flag for potential extrapolation, indicating that predictions for that landscape are less reliable because the observed spatial patterns fall outside the theoretical training domain."

3. Conceptual Distinction from Process-Based Models: As detailed in the new Section 3.4.1, we explicitly discuss the distinction between our pattern-based inference framework and traditional dynamic process-based models, which are conventionally viewed as the standard approach for incorporating physical mechanisms. While process-based models simulate the mechanistic, step-by-step physical propagation of individual events over time, our framework operates within a pattern-based statistical paradigm. It focuses on evaluating the emergent spatial attributes of disturbed landscapes. Crucially, instead of omitting these physical realities, our framework mathematically encapsulates the structural outcomes of complex spatio-temporal interactions, including linked disturbances and disturbance cascades , through a massive training library of over 8 million diverse regime combinations. This endows the model with a robust statistical generalization capability to infer long-term regimes from static spatial patterns.

RC2.5: *In a couple of places, biomass products are described as "globally consistent" but it's not clear what this means. Consistent with ground data (this isn't true)? Consistent with each other (very much not true)? Internally consistent error properties (not true)? So what is it?*

AR2.5: The term "globally consistent" was intended to mean "globally continuous" (i.e., providing wall-to-wall coverage). To avoid any ambiguity, we have replaced all phrases throughout the manuscript.

RC2.6: *Figure 2 suggests significant biases for some parameters over some ranges.*

AR2.6: Figure 2b is a density scatter plot displaying predictions across a massive dataset, visual outliers are inevitably present. However, as indicated by the logarithmic color scale representing data density, these outliers represent a mathematically negligible fraction of the total samples. The vast majority of the data points are densely clustered along the 1:1 line. This tight concentration is further supported by the high Nash-Sutcliffe Efficiency (NSE) scores (> 0.85) and the regression slopes being close to 1.0, confirming the absence of significant systematic biases across the parameter ranges**.**

RC2.7: *3.1.2 talks about aggregating 25 km tiles onto a 0.25° grid, but 0.25° corresponds to ~27.75 km (at all latitudes) so what is meant by aggregation?*

AR2.7: We clarify that 0.25° corresponds to ~27.75 km only at the equator. Due to the convergence of meridians, the physical longitudinal width of a 0.25° cell decreases significantly at higher latitudes (e.g., to approximately 13.9 km at 60° latitude). Because our 25 km × 25 km landscape tiles maintain a constant true surface area globally, the spatial relationship between these tiles and the 0.25° geographic grid varies across latitudes. In this context, "aggregation" refers to the spatial regridding process where we calculate the statistical mean (and sub-grid variance) of all independent 25 km × 25 km tiles that overlap with a given 0.25° × 0.25° grid cell. We have revised the text to explicitly describe this regridding mechanism.

CM Line 506: “Because the 25 km × 25 km tiles maintain a constant true surface area, while the physical footprint of the 0.25° grid cells varies by latitude, this spatial aggregation calculates the mean value of all independent landscape tiles that overlap with a specific 0.25° grid cell.”

RC2.8: *There is no discussion of Fig. 4 and whether the patterns shown are consistent with other estimates, such as of turnover time, fire statistics, etc.*

AR2.8: We thank the reviewer for the valuable suggestion. Conducting a comprehensive comparison and synergistic correlation analysis between our derived global patterns and other environmental variables (such as fire statistics and turnover times) is planned as the next step of our research.

RC2.9: *l. 289 - 295 This section first says uncertainty is high for the parameters, then it says it is high for β. Both can’t be true. This section seems oblivious of the fact that GlobBiomass relies on ALOS-1 data which is completely insensitive to high biomass, so what information is it supplying in these forests and how does this affect the procedure in this paper?*

AR2.9: We agree that the original phrasing was imprecise and have clarified that overall prediction uncertainty is small, though distinct regional hotspots exist for parameters like β. We respectfully argue that regarding ALOS-1, the statement that it is "completely insensitive" is incorrect. The signal has weak sensitivity to structural features and is affected by calibration uncertainty. Santoro et al. (2021) demonstrated that a combination of SAR processing, filtering, and spatially adaptive modeling can compensate for such errors. However, we agree this does not mean the AGB estimates are completely free from biases and uncertainty. In our framework, this weak sensitivity dampens the local spatial variance, which is exactly what drives the regional uncertainty hotspots for β. We have now extensively evaluated this dynamic in the newly added Section 3.4.3 Possible Implications from Potential Signal Saturation in High-Biomass Forests.

CM Line 553: “As shown in Fig. 5, the overall model-related uncertainty for all parameters is relatively small compared to their respective parameter ranges, indicating generally robust predictions globally. However, regions with relatively higher uncertainty exhibit distinct regional hotspot characteristics. This anomaly is particularly pronounced for β, with higher uncertainty consistently concentrated in the humid tropics (e.g., the Amazon and Congo basins). This suggests the model has greater difficulty disentangling the disturbance severity signal from biomass patterns in these high-biomass, structurally complex ecosystems, which is also linked to the physical saturation of satellite radar signals, a limitation that is explicitly evaluated in Section 3.4.3.”

RC2.10: *In Fig. 6 the high and low end of values are indistinguishable so it’s nearly impossible to interpret what is being shown. A better colour range is needed.*

AR2.10: Figure 6 has been replaced with an updated version featuring the improved color scale.

RC2.11: *Fig. 7: what do the colours mean? The text is too small to be legible. I have serious doubts about the authors’ interpretation of what these images say about disturbance for all three parameters, especially the high β example*

AR2.11: We replaced Figure 7 to improve legibility and expanded the explanation in Section 3.3**:**

Technical corrections

RC2.12: *Is it Kb or Kb (many discrepancies)?*

AR2.12: All instances of "Kb" have been corrected to "K_b" throughout the manuscript, figures, and supplementary materials.

RC2.13: *"Where" after an equation should be "where" in several cases.*

AR2.13: Corrected.

RC2.14: *Terminology is inconsistent: we have "landscape", "tile", "landscape tile", "landscape domain" and possibly other phrases all representing the same thing, as far as I can tell.*

AR2.14: We now exclusively use the term "landscape tile" when referring to the 25 km × 25 km spatial units to ensure clarity and consistency.

RC2.15: *The paper in several places talks about EO data but the AGB map is not EO data. It is a product derived from EO data (with other inputs); this is an important distinction.*

AR2.15: We have replaced instances of "EO data" with "EO products" or "satellite-derived biomass products" when referring to the AGB maps throughout the manuscript.

RC2.16: *The authors use the phrase "scientific plausibility" in a few places. There's nothing scientific about this; it's just plausibility.*

AR2.16: All instances of "scientific plausibility" have been corrected to "plausibility" throughout the manuscript.

RC2.17: *l.68 Le Toan*

AR2.17: Corrected.

RC2.18: *l. 145 superior skill compared to what?*

AR2.18: We have revised the text **CM Line 299:** "This data-driven product is constrained by in-situ eddy covariance observations and provides a reliable representation of spatial variation in GPP."

RC2.19: *In the supplement, table S2: what is pathes? Below, "integer"*

AR2.19: We corrected the spelling errors in Table S2. "Pathes" has been corrected to "patches", and "integrer" to "integer".

RC2.20: *S1.3. L_d is given 2 different definitions*

AR2.20: We corrected the notation error in Section S1.3. Continuous background mortality is now correctly denoted as L_b instead of mistakenly repeating L_d

RC2.21: *Next para: what are the initial conditions and what effect do they have?*

AR2.21: We clarified that simulations start from bare ground (zero biomass). Because the model runs for 200 years to reach a dynamic equilibrium (a shifting mosaic steady state), the initial conditions are completely overwritten and have no effect on the final extracted statistics.

RC2.22: *How can a suite of metrics be methodogically identical to anything?*

AR2.22: We have rephrased the text to state that the extraction methodology is identical, rather than the metrics themselves.

RC2.23: *Fig S2.2 Labelling illegible*

AR2.23: We have updated Figure S2.2 layout, significantly increasing the font size and legibility.

RC2.24: *S2.2 This talks about influential features but the meaning of this is never explained. Influential on what? How is influence measured? Do we really need 17 metrics?*

AR2.24: We clarified that "influence" strictly refers to the Feature Importance scores derived from the Random Forest model.

RC2.25: *Fig S2.3 What's a simulation-observation gap?*

AR2.25: We replaced the vague term "simulation-observation gap" with "statistical discrepancy between simulated and observed biomass features".

RC2.26: *S3 is full of confusing labelling, sometimes f, sometimes j, and no statement that x is a vector*

AR2.26: We updated the mathematical notation in S3, including training sample vectors ($\{x_i\}_{i=1}^{n}$) and unified all feature indices to j.

RC2.27: *S3.2 talks about Section 2.5.1; there is no such section*

AR2.27: Corrected by Section S3.1.

RC2.28: *Fig S4.1 conveys almost no information as it all just looks bluey-grey; it needs a much better choice of display format.*

AR2.28: We have updated Figure S4.1 with a high-contrast, discrete color palette.