



## RESEARCH ARTICLE

10.1029/2023MS004099

# Understanding Disturbance Regimes From Patterns in Modeled Forest Biomass

 Siyuan Wang<sup>1,2</sup> , Hui Yang<sup>1</sup>, Sujan Koirala<sup>1</sup> , Matthias Forkel<sup>2</sup> , Markus Reichstein<sup>1,3</sup> , and Nuno Carvalhais<sup>1,3,4</sup> 

<sup>1</sup>Max-Planck Institute for Biogeochemistry, Jena, Germany, <sup>2</sup>Technische Universität Dresden, Institute of Photogrammetry and Remote Sensing, Dresden, Germany, <sup>3</sup>ELLIS Unit Jena at Michael-Stifel-Center Jena for Data-driven and Simulation Science, Jena, Germany, <sup>4</sup>Departamento de Ciências e Engenharia do Ambiente, Faculdade de Ciências e Tecnologia, Universidade Nova de Lisboa, Caparica, Portugal

**Key Points:**

- We investigate the link between disturbance regimes and spatial patterns of aboveground biomass emerging from diverse primary productivity
- The proposed framework allows for inferring disturbance probability, size and intensity from spatial features in aboveground biomass
- Disturbance regimes from high-res Earth observations can enhance carbon cycle dynamics prediction from inter-annual to longer time scales

**Supporting Information:**

Supporting Information may be found in the online version of this article.

**Correspondence to:**

S. Wang and N. Carvalhais,  
swang@bgc-jena.mpg.de;  
ncarvalhais@bgc-jena.mpg.de

**Citation:**

Wang, S., Yang, H., Koirala, S., Forkel, M., Reichstein, M., & Carvalhais, N. (2024). Understanding disturbance regimes from patterns in modeled forest biomass. *Journal of Advances in Modeling Earth Systems*, 16, e2023MS004099. <https://doi.org/10.1029/2023MS004099>

Received 31 OCT 2023

Accepted 20 MAY 2024

**Author Contributions:**

**Conceptualization:** Siyuan Wang, Nuno Carvalhais

**Data curation:** Siyuan Wang

**Formal analysis:** Siyuan Wang

**Funding acquisition:** Markus Reichstein, Nuno Carvalhais

**Investigation:** Siyuan Wang, Hui Yang, Nuno Carvalhais

**Methodology:** Siyuan Wang, Hui Yang, Sujan Koirala, Nuno Carvalhais

**Abstract** Natural and anthropogenic disturbances are important drivers of tree mortality, shaping the structure, composition, and biomass distribution of forest ecosystems. Differences in disturbance regimes, characterized by the frequency, extent, and intensity of disturbance events, result in structurally different landscapes. In this study, we design a model-based experiment to investigate the links between disturbance regimes and spatial biomass patterns. First, the effects of disturbance events on biomass patterns are simulated using a simple dynamic carbon cycle model based on different disturbance regime attributes, which are characterized via three parameters:  $\mu$  (probability scale),  $\alpha$  (clustering degree), and  $\beta$  (intensity slope). 856,800 dynamically stable biomass patterns were then simulated using combined disturbance regime, primary productivity, and background mortality. As independent variables, we use biomass synthesis statistics from simulated biomass patterns to retrieve three disturbance regime parameters. Results show confident inversion of all three “true” disturbance parameters, with Nash-Sutcliffe efficiency of 94.8% for  $\mu$ , 94.9% for  $\alpha$ , and 97.1% for  $\beta$ . Biomass histogram statistics primarily dominate the prediction of  $\mu$  and  $\beta$ , while texture features have a more substantial influence on  $\alpha$ . Overall, these results demonstrate the association between biomass patterns and disturbance regimes. Given the increasing availability of Earth observation of biomass, our findings open a new avenue to understand better and parameterize disturbance regimes and their links with vegetation dynamics under climate change. Ultimately, at a large scale, this approach would improve our current understanding of controls and feedback at the biosphere-atmosphere interface in the present Earth system models.

**Plain Language Summary** Forest dynamics are shaped by different disturbances, which are challenging to monitor and predict. Identifying individual disturbance occurrences and their impact on forest carbon stocks (biomass) is complex. However, our study deciphers the characteristics of disturbance occurrence, that is, disturbance regime, from biomass pattern. We characterized this regime across three dimensions: extent ( $\mu$ ), frequency ( $\alpha$ ), and intensity ( $\beta$ ). Through a 200-year landscape experiment, we explored the synthetic dynamically stable biomass under different disturbance regimes. Statistical features from biomass simulations revealed distinct spatial patterns, forming a connection between these patterns and the disturbance regime parameters via machine learning. Notably, specific biomass pattern statistics influence distinct disturbance regime parameters:  $\mu$  and  $\beta$  are linked to histogram stats, while  $\alpha$  is tied to texture statistics. This approach establishes a framework to diagnose disturbance regimes from biomass patterns, offering a way to incorporate these regimes into Earth system models.

## 1. Introduction

Mortality is one of the key processes of vegetation dynamics (Franklin et al., 1987; Runkle, 2000) that dominates aboveground carbon turnover (Carvalhais et al., 2014; Thurner et al., 2016) and contributes to model uncertainties in the carbon cycle (Friend et al., 2014). The diverse range of natural (e.g., fires, droughts, wind-throws, pathogens and insects outbreaks) and anthropogenic disturbances (e.g., agricultural expansion, urbanization, and clearcutting) act as strong drivers of vegetation mortality, leading to the total or partial loss of biomass (Grime, 1977; Hammond et al., 2022; McDowell et al., 2022). A better understanding of mortality and disturbance and their impacts on carbon dynamics is thus crucial for constraining future carbon cycling prognostics (Friend et al., 2014).

© 2024 The Authors. Journal of Advances in Modeling Earth Systems published by Wiley Periodicals LLC on behalf of American Geophysical Union. This is an open access article under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

**Supervision:** Hui Yang, Sujan Koirala, Matthias Forkel, Markus Reichstein, Nuno Carvalhais

**Visualization:** Siyuan Wang

**Writing – original draft:** Siyuan Wang, Hui Yang, Nuno Carvalhais

**Writing – review & editing:**

Siyuan Wang, Hui Yang, Sujan Koirala, Matthias Forkel, Markus Reichstein, Nuno Carvalhais

The mortality caused by disturbances, as well as primary productivity and allocation, play an important role in controlling the local distribution and the large scale spatial gradients of above-ground biomass (AGB, Delbart et al., 2010; Johnson et al., 2016). However, varying levels of mortality, productivity and disturbance events may lead to similar biomass levels, presenting an equifinality issue (Ryan & Williams, 2011; Williams et al., 2013). This challenge, combined with the stochastic nature of disturbances, continues to hinder the understanding of disturbance regimes relate to biomass patterns (Allen et al., 2010; Chambers et al., 2004; Hammond et al., 2022). Characteristics of disturbances at the landscape level, that is, disturbance regimes, are commonly inferred using metrics like size, frequency, intensity, and aggregation (Turner, 2010), which describe the cumulative effects of all disturbance events in a given area and time period (Senf & Seidl, 2021a). The disturbance regime ultimately leads to a shifting steady-state mosaic, represented by patches of distinct successional stages or carbon stocks over long time periods (Brokaw & Scheiner, 1989).

Most research on quantifying disturbance regime parameters has been carried out either through observation-driven methods or by using model-data-integration. For example, remote sensing based on spectral bands, indices and segment outlines have been used to identify changes in vegetation that are associated with disturbance magnitude, duration, and rate of change (Chambers et al., 2013; Senf & Seidl, 2021a). Williams et al. (2013) utilized satellite-based biomass observations to determine two disturbance parameters: average probability and intensity; realizing a high level of equifinality when analyzing total biomass. Additionally, by analyzing consecutive biomass maps, several studies have been able to detect differences in patterns of intensity, ranging from deforestation to widespread low-intensity disturbance (Hill et al., 2015). Moreover, the clustering pattern of disturbance events has been recognized as a distinguishing attribute of different disturbance regimes and failure to resolve these patterns can lead to misestimation of average mortality and growth patterns (Fisher et al., 2008). These studies inspired the subsequent investigation, and highlighted the significant potential of using biomass observations to understand and quantify disturbance regimes. In parallel, the emergence of up-to-date biomass observations (Saatchi et al., 2011; Santoro et al., 2021, 2022) opens novel pathways for comprehending and investigating the clustering patterns of disturbances, alongside other facets of the disturbance regimes, at a more intricate scale. Ultimately, disturbance parameters based on observations could be incorporated in process-based models (e.g., Friend et al., 2014) or individual-based models (Bossel & Krieger, 1994; Bugmann, 2001; Köhler & Huth, 1998; Yan et al., 2005) as stochastic model components that allow the quantification of the impacts of disturbances on the forest carbon cycle from local and global scales.

The goal of our study is to comprehensively characterize disturbance regimes and investigate their connection with resulting biomass patterns through a model-based experiment. Specifically, we focus on the methodology used to simulate the impact of three distinct disturbance regime attributes - extent, frequency, and intensity of disturbance events - on biomass dynamics. Below, in Section 2, we introduce in detail the carbon model, the parameters that control the simulated disturbance regimes, describe the approach to generate disturbance cubes, and the methodology for retrieving disturbance parameters. In Section 3, we present the results of these experiments, namely on the varying patterns of biomass emerging from different disturbance regimes, and the performance of the regression approach to invert the underlying disturbance parameters. This is followed by a discussion on the robustness of the framework and an outlook in Section 4. We present the conclusions of our research in Section 5.

## 2. Methods

### 2.1. Dynamic Carbon Cycle Model

We simulate the carbon cycle at the patch level. Each patch is specified as a homogeneous forest stand, representing the smallest computing unit during the simulation (Fisher et al., 2008). The changes in aboveground vegetation carbon ( $C$ , in  $kgC\ m^{-2}\ yr^{-1}$ ) are determined by the difference between aboveground carbon gains (via photosynthesis,  $NPP_{AGB}$ ) and losses ( $L$ , Equation 1) at the annual scale as,

$$\frac{dAGB}{dt} = NPP_{AGB} - L \quad (1)$$

The aboveground carbon gain,  $NPP_{AGB}$ , is calculated from gross primary production (GPP), the losses of C to growth respiration ( $1 - Y_G$ , Amthor, 2000), and the fraction of biomass that is aboveground ( $f_{AGB}$ ) as,

**Table 1**  
Parameter Setting

Parameter	Definition	Range	Interval	Count
$\mu$	Probability scale	[0.01:0.05]	0.005	9
$\alpha$	Clustering degree	[1.0:1.8]	0.05	17
$\beta$	Intensity slope	[0.03:0.5]	0.01/0.05/0.1	14
$G$	Photosynthetic capacity	[0.03:0.1]	0.01/0.02/0.03	5
$K_b$	Background mortality	[0.025:0.2]	0.025	8

$$NPP_{AGB} = GPP \times (1 - Y_G) \times f_{AGB} \quad (2)$$

To simplify the experiments, the transfer ratio from  $GPP$  to  $NPP_{AGB}$  ( $(1 - Y_G) \times f_{AGB}$ ) is fixed to a value of 0.5, representing 2/3 of C allocation to AGB and a growth respiration ratio of 0.25 (Amthor, 2000).

The GPP dynamics are represented as a simple saturating exponential function of biomass. We acknowledge that the variations in the relationship between GPP and AGB for which we introduce a varying parameter  $G$  (Equation 3). We fixed two constants in Equation 3, A and B, set to 100 and 1,200, respectively, to establish a realist optimal biomass growth curve per

patch reaching an asymptotic maximum with age. The parameter  $G$  is left as a free parameter to regulate the maximum photosynthetic capacity, hence the maximum biomass and the recovery rate simultaneously, representing the impact of species and local edaphoclimatic conditions in primary productivity dynamics between different landscapes.

$$GPP = \frac{A}{G + e^{-\frac{AGB}{B}}} \quad (3)$$

The total carbon loss includes the carbon losses during disturbance events ( $L_d$ ) and by the background mortality ( $L_b$ , Equation 4) as,

$$L = L_b + L_d \quad (4)$$

The  $L_b$  is assumed to be a constant proportion of  $AGB$ , implicitly including the average effects of litterfall, root exudates, and herbivory (Thurner et al., 2016) as,

$$L_b = AGB \times K_b \quad (5)$$

$$K_b = \frac{1}{\tau} \quad (6)$$

where  $K_b$  refers to the background mortality rate, a reciprocal of turnover time ( $\tau$ ). To account for the fact that  $K_b$  is also spatially variable (e.g., Thurner et al., 2016), we define a range of  $K_b$  between 0.025 and 0.2, with an interval of 0.025, representing background turnover times from 5 to 40 years (Table 1). The other part of the carbon loss is caused by disturbance,  $L_d$ , determined by the intensity of disturbance event covering the patch as,

$$L_d = AGB \times f_L \quad (7)$$

where, parameter  $f_L$  represents fraction of carbon loss during the disturbance, and it depends on the size of the event and the intensity slope ( $\beta$ , see Section 2.2).

## 2.2. Modeling Different Disturbance Regimes

We applied three parameters to describe different disturbance regimes: the probability of disturbances,  $\mu$ ; the clustering patterns of events,  $\alpha$ ; and the intensity slope,  $\beta$ . These parameters represent the fraction of the domain affected by disturbances, the number and size of disturbed clusters of patches, and the fraction of carbon loss during each event, respectively. For distributing a sufficiently large and spatially random number of disturbance events, our experiment is set on a domain of 1,000 by 1,000 pixels to simulate the corresponding landscape-scale domains. As in Fisher et al. (2008), here we assume that the smallest disturbance event size is at the patch level, with a roughly specified canopy area of approximately  $10 \times 10 \text{ m}^2$ , resulting in a domain size of circa  $100 \text{ km}^2$ .

### 2.2.1. Parameterization of Disturbance Regimes

Parameter  $\mu$  refers to the total disturbed fraction of the domain, where  $D$  refers to the total domain size and  $D_a$  is the area of the domain affected by disturbances as,

$$D_a = D \times \mu \quad (8)$$

We followed Fisher's method by categorizing events into 20 size classes and assigning them specific quantities using parameter  $\alpha$ , which was originally defined as the clustering degree and is the scaling exponent used to determine the number of events of each particular size (Fisher et al., 2008):

$$n_{z_i} = A z_i^{-\alpha} \quad (9)$$

where  $z_i$  is the size of the specific class,  $\alpha$  is the scaling exponent,  $n_{z_i}$  represents the number of disturbance events at a specific event size.  $A$  is the proportionality factor, manipulated by the size of the total disturbed area and the setting of events size series as,

$$A = \frac{D_a}{\sum (z_i \cdot z_i^{-\alpha})} \quad (10)$$

where  $z_i$  is the corresponding disturbance event size. We applied an event size ranging between a patch to half the size of the domain during the simulations (see details in Text S1 of the Supporting Information S1).

For parameter  $\beta$ , we assumed that the intensity of disturbance ( $f_L$ , fraction of carbon loss per event) is proportional to the logarithm of event size ( $\log_{10}(Z_i)$ , Chambers et al., 2013). This relationship between intensity and event size is controlled by parameter  $\beta$  (intensity slope, Figure 3), and  $b$  is a constant value:

$$f_L = \beta \cdot \log_{10}(z_i) + b \quad (11)$$

More details on the parameterization setup related to  $\alpha$  and  $\beta$  (Equations 9–11) can be found in Text S1 of the Supporting Information S1.

### 2.2.2. Disturbance Parameter Ranges

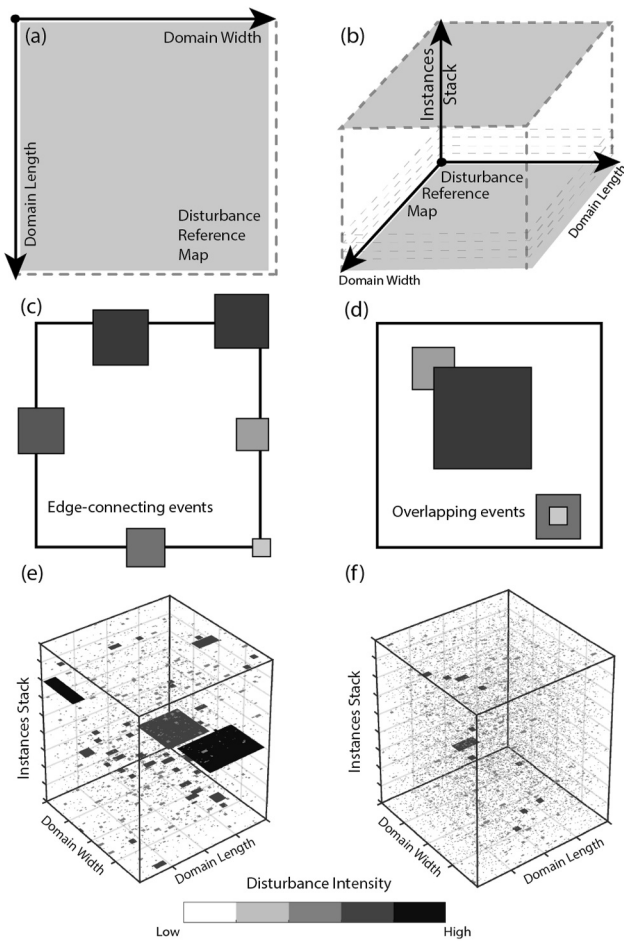
The inclusion of the parameters  $\mu$ ,  $\alpha$ , and  $\beta$  allows a flexible description of disturbance regimes in different landscapes, containing a wide range of all possible disturbance regimes without specifying disturbance type. For providing realistic simulation scenarios we consulted existing literature to set different ranges and intervals for the primary productivity parameter  $G$ , background mortality rate  $K_b$ , and the three disturbance parameters independently (Table 1).

In particular, the parameter  $G$  is specified as a range from 0.03 to 0.1, imposing a range in GPP of 1 and 4  $kgC m^{-2} yr^{-1}$  and an average dynamically stable biomass from 10 to 40  $kgC m^{-2}$ , without disturbances. The parameter  $\mu$  is set in a range of 0.01–0.05 with an interval of 0.005; this range substantially spans the average value of 0.02 documented in forests by Moorcroft et al. (2001) and Malhi et al. (2004). The settings for the clustering parameter  $\alpha$  comes from observed gap-size distributions from previous studies in the tropical and sub-humid forest ecosystem, which indicated a range between 1.1 and 1.6 (Araujo et al., 2021; Fisher et al., 2008; Jans et al., 1993; Lawton & Putz, 1988; Nelson et al., 1994; Yavitt et al., 1995). In the experiment, we have increased the documented range slightly from 1.0 to 1.8, with an interval of 0.05. For parameter  $\beta$ , we have considered the findings from Chambers et al. (2013) to establish a range between 0.03 and 0.5. The intervals within that range differ, with a value of 0.01 for the range [0.03, 0.09], 0.05 for [0.1, 0.25], and 0.1 for [0.3, 0.5] (see Figure 3).

Given the current parameter ranges and the simulation length of 200 years, the current setup is bound to disturbance regimes occurring no less that once every 100 years (minimum  $\mu$  of 0.01) for the given domain size. Capturing very low frequency events would necessarily require expanding the simulation length, ensuring the emergence of a dynamically stable biomass distribution dependent on extremely rare disturbance events.

### 2.3. Generation of Disturbance Events

A stochastic disturbance event generator was designed to prescribe disturbances in a 1,000 by 1,000 simulation grid (the simulation domain) according to different disturbance regime parameters that control the size and intensity of each single event. We implemented Fisher et al. (2008)'s approach for distributing disturbance events of



**Figure 1.** Conceptual diagram of disturbance reference cubes, (a) displays a two-dimensional disturbance reference map, which represents a snapshot composed of a  $1,000 \times 1,000$  array from a disturbance cube array, (b) showcases a three-dimensional disturbance reference cube consisting of 200 snapshots. Each snapshot simulates a unique stochastic spatial distribution of disturbance events, and the cube features a distinct combination of  $\mu$  and  $\alpha$ . It is essential to note that the disturbance events within the cube should not (c) interface with edges or (d) overlap with one another, (e) and (f) depict two examples of disturbance reference cubes featuring different disturbance regimes.

statistics from the GLCMs method, namely: contrast; correlation; energy; and homogeneity (Table 2). For doing so, all biomass maps were re-scaled to a range of 0–255, and AGB outliers were substituted with the nearest neighboring pixel prior to rescaling to avoid contamination in texture features. These outliers are individual isolated pixels with substantially higher AGB values resulting from incidentally undisturbed locations during the 200 years of simulations.

The simulation of biomass dynamics using the carbon cycling model involved a comprehensive examination of 85,680 parameter combinations. This comprised of 2,142 disturbance regimes in combination with 5 primary productivity scenarios and 8 background mortality rate scenarios. Each parameter combination was executed 10 times, with the reference disturbance cube being shuffled in a different sequence for each run. Ultimately, 17 statistics were extracted for each run, including mean GPP at the end of simulation, as well as 16 other statistics related to the steady-state biomass distribution.

different sizes per year. For each year the disturbance events (in rectangular and discussed in Section 4.1.2) are randomly placed in the domain by size in descending order, without overlapping each other or exceeding the domain limits. In this way, the generator ensures an accurate prescription of each disturbance regime for each simulation.

A full factorial combination of the different parameter's ranges and intervals sums up to a total of 2,142 combinations of disturbance regimes. For every disturbance regime, we generate a reference disturbance data set for a  $1,000 \times 1,000$  domain per year, for a period of 200 years, in the form of a three-dimensional array (a  $1,000 \times 1,000 \times 200$  data cube) (Figures 1e and 1f). As such, every snapshot of the reference cube is a  $1,000 \times 1,000$  stochastic disturbance reference map, providing the spatial reference for simulating the effect of disturbances in carbon cycling dynamics. To further impose variability in the resulting biomass distributions, for the same disturbance regime, the disturbance data set was shuffled 10 times in the time domain, generating a total of more than  $2 \times 10^4$  simulation scenarios. The emerging AGB distributions from the different scenarios result from the variations in disturbance regimes, in productivity ( $G$ ), and in background mortality ( $K_b$ ) levels (see details in Text S2 of the Supporting Information S1), upon which three types of statistic metrics were used to characterize the biomass patterns at steady state on the domain scale.

#### 2.4. Characterizing the Biomass Patterns

The first type of statistics utilized in our study is based on the histogram distribution, including mean, median, variance, skewness, kurtosis, percentiles, as well as standard deviation and coefficient of variation. Previous literature suggests that some of these metrics, such as skewness and mean, are associated with the probability and intensity of disturbances (Williams et al., 2013). We have additionally introduced information theory based metric: the Shannon entropy index (Shannon, 1948), also called Shannon-Wiener index, which is a widely used indicator for describing the diversity level in ecosystems (Spellerberg & Fedor, 2003). Furthermore, we included statistical properties based on the texture of biomass distribution. Texture provides information about the spatial arrangement of intensities in an image (e.g., continuity, contrast, smoothness), in our case, the emergent biomass map. We utilized Gray-Level Co-Occurrence Matrices (GLCMs), one of the most common texture feature extraction methods based on image statistics, to study the spatial correlation properties by using gray-scaled images (Haralick et al., 1973). To characterize texture on the biomass maps we applied four

**Table 2**  
*Statistics of the Steady-State Biomass Map*

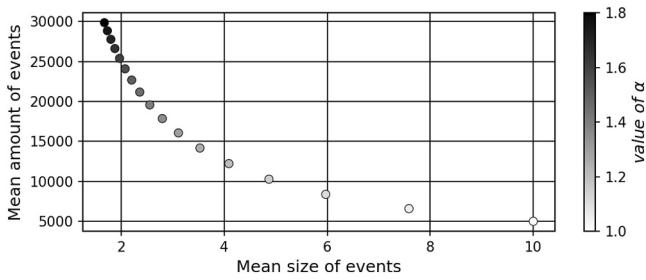
Feature types	Biomass statistics	Variable names	Formula	
Histogram features	Mean	AGBmean	$\frac{\sum_{i=1}^N A_i}{N}$	N: Total patch amount $A_i$ : Biomass value for patch $i$
	Median	AGBmed	$Med(A)$	A: Biomass map
	Range	AGBrange	$P_{90} - P_{10}$	P90: Percentile 90% P10: Percentile 10%
	Variance	AGBvar	$\frac{1}{N-1} \sum_{i=1}^N  A_i - \mu ^2$	$\mu$ : Mean biomass
	standard deviation	AGBstd	$\sqrt{\frac{1}{N-1} \sum_{i=1}^N  A_i - \mu ^2}$	
	coefficient of variation	AGBcv	$100 \times \frac{\sigma}{\mu}$	$\sigma$ : Standard deviation
	Skewness	AGBskew	$\frac{\sum_{i=1}^N (A_i - \mu)^3}{(N-1)\sigma^3}$	
	Kurtosis	AGBkurt	$\frac{\sum_{i=1}^N (A_i - \mu)^4}{(N-1)\sigma^4}$	
	Percentile 25%	AGBp25	$P_{25}$	P25: Percentile 25%
	percentile 75%	AGBp75	$P_{75}$	P75: Percentile 75%
	Trimean	AGBtrim	$(P_{25} + 2 \times MED + P_{75})/4$	MED: Median value
Informative feature	Shannon entropy	Shannon	$-\sum p \cdot \log_2(p)$	$p$ : Normalized histogram counts
Texture features	Contrast	Contrast	$\sum_{i,j}  i-j ^2 glc(i,j)$	$i$ : Reference pixel value $j$ : Neighbor pixel value
	Correlation	Correlation	$\sum_{i,j} \frac{(i-\mu_i)(j-\mu_j) glc(i,j)}{\sigma_i \sigma_j}$	$glc(i,j)$ : An entry in GLCM $\mu_i, \mu_j$ : Means of GLCM w.r.t. $i$ and $j$
	Energy	Energy	$\sum_{i,j} glc(i,j)^2$	$\sigma_i, \sigma_j$ : Standard deviations of GLCM w.r.t. $i$ and $j$
	Homogeneity	Homogeneity	$\sum_{i,j} \frac{glc(i,j)}{1+ i-j }$	

## 2.5. Prediction and Validation

We used the outputs from the dynamic model for the last 10 simulation years, when there is a dynamically stable biomass distribution, to compute an average biomass per grid cell. The resulting biomass matrix is used to estimate the landscape biomass metrics (described in the above Section 2.4). These landscape metrics are the inputs to train a multi-output random forest model for predicting  $\mu$ ,  $\alpha$ , and  $\beta$ .

We used a multi-output random forest regression method in Scikit-learn (Pedregosa et al., 2011) to investigate the relationship between the simulated biomass statistics and the prescribed disturbance regimes. Cross-validation methods are crucial in mitigating overfitting and enhancing generalization in machine learning models. These methods assess the robustness of a model by systematically excluding a portion of data from training and evaluating it across various data subsets. To evaluate the generalization ability of models the cross-validation follows three different strategies: Completely Random 10-fold method (CR), the Leave-One-Sequence-Out method (LOSO), and the Leave-One-Parameter-Out method (LOPO). The cross-validation tests additional to the random 10-fold aim to evaluate the strength of the approach to predict unseen parameters (LOSO) or unseen complete disturbance regimes (LOPO) during the training phase, more details in Text S3 in Supporting Information S1. Extrapolation in the random forest model leads to unreliable or inaccurate predictions when confronted with data that deviates significantly from the training data. We have devised a supplementary cross-validation test called LOPO + LOSO to address this problem.

We use the Nash-Sutcliffe efficiency (NSE) to evaluate the performance of our prediction model (Nash & Sutcliffe, 1970). The NSE measures the accuracy of the predicted disturbance regime parameters compared to the prescribed values in this study. A higher NSE indicates better accuracy. The formula for calculating NSE (D. N. Moriasi et al., 2007) is shown in Equation 12,



**Figure 2.** The parameter  $\alpha$  exerts control over the total number and average sizes of events in all disturbance cubes generated with the same total disturbed area (i.e., same  $\mu$ ), exhibiting an exponential relationship. Specifically, an increase in  $\alpha$  results in a higher proportion of relatively small events, whereas a decrease in  $\alpha$  tends to produce fewer, yet larger events.

$$NSE = 1 - \frac{\sum_{i=1}^n (Y_i^{obs} - Y_i^{pre})^2}{\sum_{i=1}^n (Y_i^{obs} - \bar{Y}^{obs})^2} \quad (12)$$

where the  $Y_i^{obs}$  is the  $i$ th prescribed disturbance regime parameter,  $Y_i^{pre}$  is the  $i$ th predicted value for the corresponding parameter,  $\bar{Y}^{obs}$  is the mean of the prescribed parameter, and  $n$  is the total number of observations.

The importance of a feature is estimated as the (normalized) reduction in prediction error brought by the inclusion of that feature in the RF model (Pedregosa et al., 2011). The value of feature importance is directly proportional to its contribution for predicting the target variables, in this case, for predicting disturbance regime parameters.

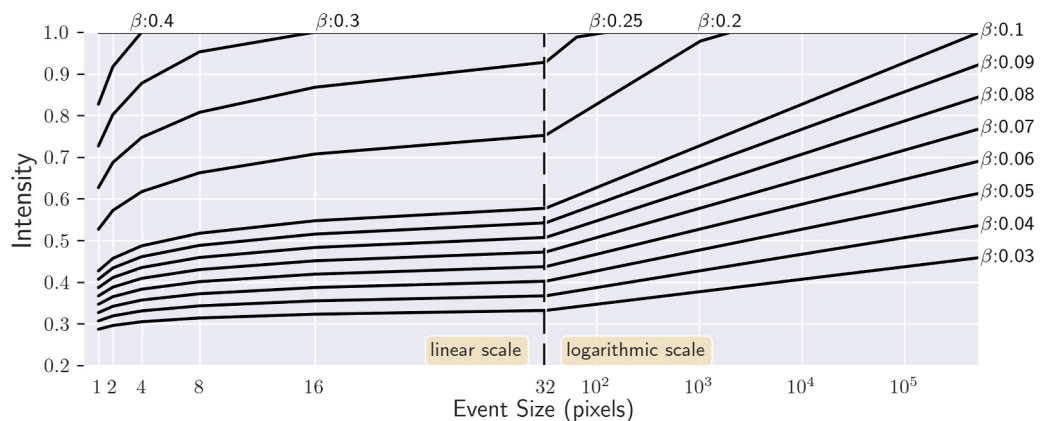
### 3. Results

#### 3.1. Parameterization and Dynamic Carbon Model

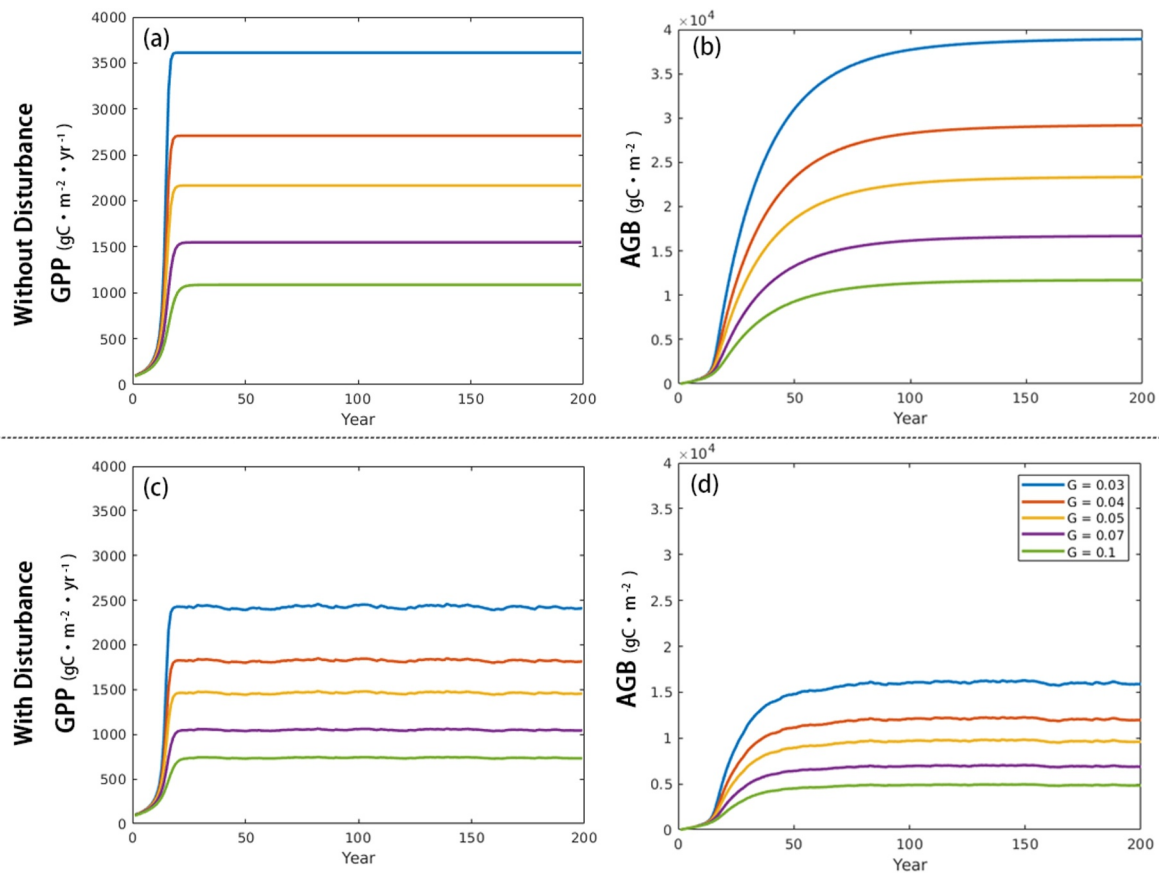
##### 3.1.1. Parameterization of Disturbance Regimes

The clustering parameter  $\alpha$  defines the relationship between the size and frequency of disturbance events across the domain establishing that the number of events decreases exponentially with the event size (see details in Figure S1 of the Supporting Information S1). This relationship is diagnosed from the simulated cubes. The average size of events in the domain exponentially decays as the parameter  $\alpha$  increases, conversely, the number of events tends to a logistic increase, which confirms that larger  $\alpha$  corresponds to more and smaller discrete events rather than few and larger ones. With discrete but progressive  $\alpha$  values, the disturbance data cubes provide a gradient in the relationship between the amounts and sizes of events across the domain (Figure 2).

We have introduced a range in the event intensity parameter,  $\beta$ , that results in an intensity gradient for events of the same size between the disturbance data cubes. The relationship between  $\beta$  and event size, as defined in Equation 11, results in a logarithmic increase in disturbance intensity as the event size increases. Figure 3 shows this relationship for all the assigned  $\beta$  values in our experiment. The results demonstrate the distinct gradient across all curves, each translating a steeper increase between intensity and event sizes and  $\beta$  increases. When  $\beta$  was greater than 0.2, the disturbance intensity was 100% for events larger than  $\sim 2 \times 10^3$ . Notably,  $\beta$  values of 0.5 impose a full intensity disturbance (full loss) for events of any size (Figure 3).



**Figure 3.** Logistic correlation between the intensity of disturbance event and its size. The Y-axis displays the intensity value, which refers to the proportion of biomass loss attributed to the event, while the X-axis represents the gradient event size. The range of sizes is divided into two parts for ease of visualization: a linear scale for events under 32 pixels and a logarithmic scale for larger events. Notably, the curve of  $\beta=0.5$  is indistinct since it saturates at an intensity of 1 from the outset of the event size.



**Figure 4.** The evolution of above-ground biomass and gross primary production is examined under varying values of Parameter  $G$ . Figures (a) and (b) represent the scenario without any disturbance events, while figures (c) and (d) demonstrate the impact of a disturbance regime with  $\mu=0.03$ ,  $\alpha=1.0$ , and  $\beta=0.2$ .

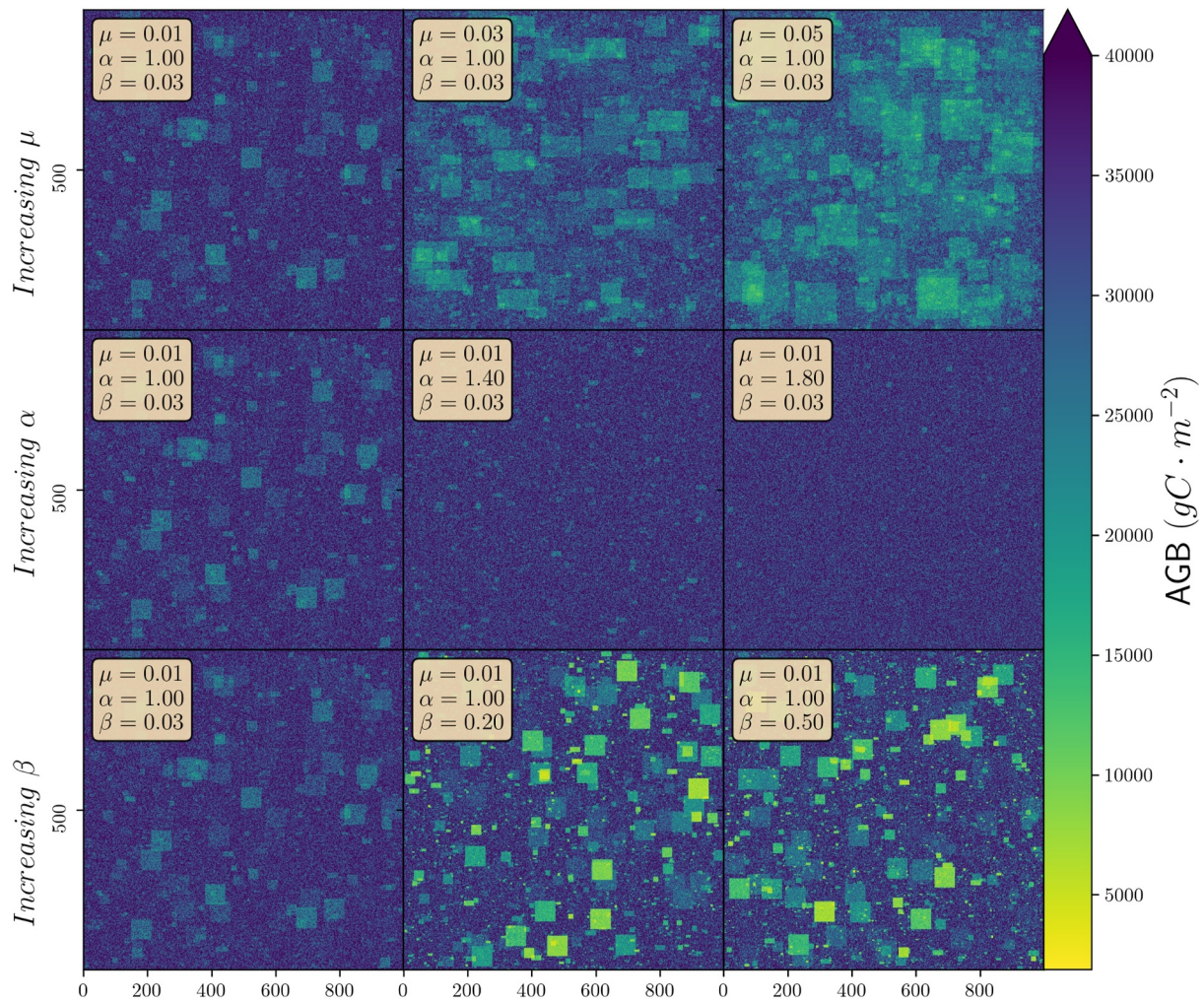
### 3.1.2. Temporal Carbon Dynamics

By employing the parsimonious carbon model, we can analyze the dynamics of AGB given trends in GPP alongside the carbon losses due to the background mortality and to the disturbance regime (Figure 4). The range of values for parameter  $G$ , designed to represent variations in the relationship between photosynthesis and biomass, produce a gradient in maximum GPP at steady states, with and without taking disturbances into account. Such gradient directly compares to the gradients in maximum AGB, both across the GPP gradients driven by  $G$ , with noteworthy variations in average biomass across domains (from  $\sim 10 \text{ kgC m}^{-2}$  to  $40 \text{ kgC m}^{-2}$ ) under a no disturbance scenario. When considering a disturbance regime with parameters  $\mu = 0.03$ ,  $\alpha = 1.0$ , and  $\beta = 0.2$ , the steady-state average levels of GPP and AGB clearly exhibit a decrease. Yet, and particularly for this disturbance regime, the reduction in AGB ( $\sim 58\%$ ) was more than twice the reductions in GPP ( $\sim 23\%$ ). Additionally, with the introduction of disturbance events, the original smooth growth curves are replaced by higher frequency fluctuations that become increasingly pronounced with higher values of  $G$ . The comparison of GPP and AGB evolution over time in different disturbance regimes is provided in Figure S2 of the Supporting Information S1.

### 3.1.3. Steady-State Biomass Distribution

Under the same photosynthetic capacity (fixed  $G$ ) and background mortality rate (fixed  $K_b$ ), the dynamically stable biomass (average level of biomass of the last 10 yr in the simulation) exhibits diverse spatial patterns under different disturbance regimes, prescribed according to the total disturbance extent parameter  $\mu$ , clustering degree parameter  $\alpha$ , and intensity indicator parameter  $\beta$ . The combination of the three disturbance regime parameters resulted in a large diversity of spatial patterns and biomass distributions as shown in Figure 5. Increasing parameter  $\mu$  while keeping the other two disturbance regime parameters fixed (first row in Figure 5) imposes an increasing fraction of the domain affected by disturbance that leads to more areas with lower biomass. As the





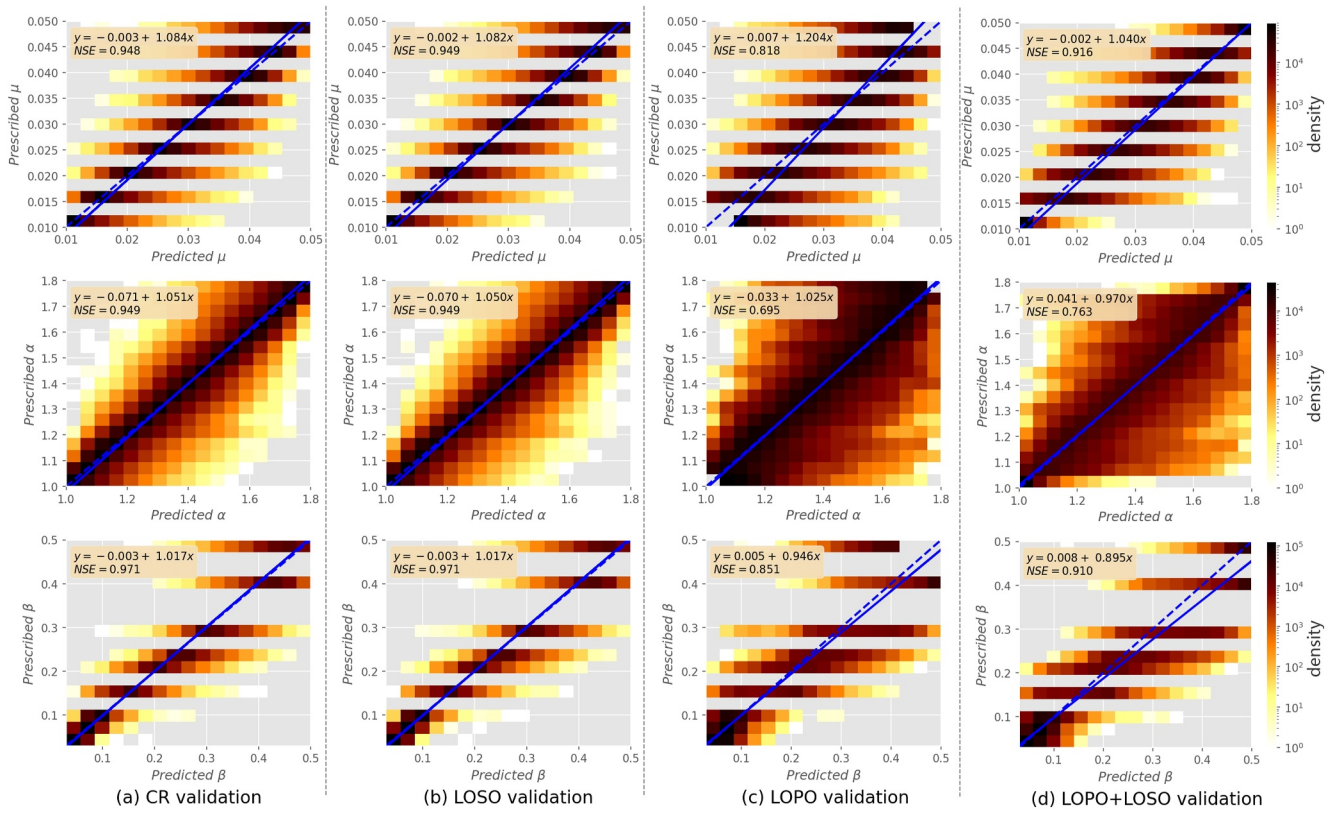
**Figure 5.** The map of steady-state biomass is compared across different disturbance regimes. The first row illustrates the impact of increasing  $\mu$ , which results in more areas with low biomass values. In the second row, the effect of increasing  $\alpha$  is represented, and the low biomass areas tend to be more scattered rather than clustered in big events. In the third row, the consequences of increasing  $\beta$  are demonstrated, with more pronounced “prints” left by disturbances.

clustering parameter  $\alpha$  increases (second row in Figure 5), the same fraction of disturbance (same  $\mu$ ) is imposed through a larger number of small events. As such, the spatial distribution of the biomass becomes more homogeneous as  $\alpha$  increases, shifting from distinguishable disturbance effects to a uniformly distributed noisy mosaic (from left to right, second row in Figure 5). The impact of higher  $\beta$  values is also evident, with more intense events resulting in greater biomass removal and lower plant regrowth levels, resulting in more contrasting disturbance footprints. The combination of the three disturbance regime parameters resulted in a large diversity of spatial patterns and biomass distributions, ideal to explore the retrieval of disturbance regimes from potential AGB observations via dynamic vegetation modeling.

### 3.2. Retrieving Disturbance Regimes From Biomass Patterns

#### 3.2.1. Cross-Validation

Figure 6 shows that all the strategies of cross-validation have a good performance for retrieving the three disturbance regime parameters under various primary productivity and background mortality, namely when evaluating on: holdout random data cubes from training (CR); holdout full disturbance regimes (full parameter vectors, LOSO); holdout parameter values (LOPO); and testing for extrapolation (LOPO + LOSO). Relying on the 16 statistics (Table 2) calculated from the biomass and the average GPP of the last year of the simulation, all the NSEs of  $\mu$ ,  $\alpha$ , and  $\beta$  exceeded 0.94 in the CR and LOSO validation. In addition, the LOPO validation exhibits a



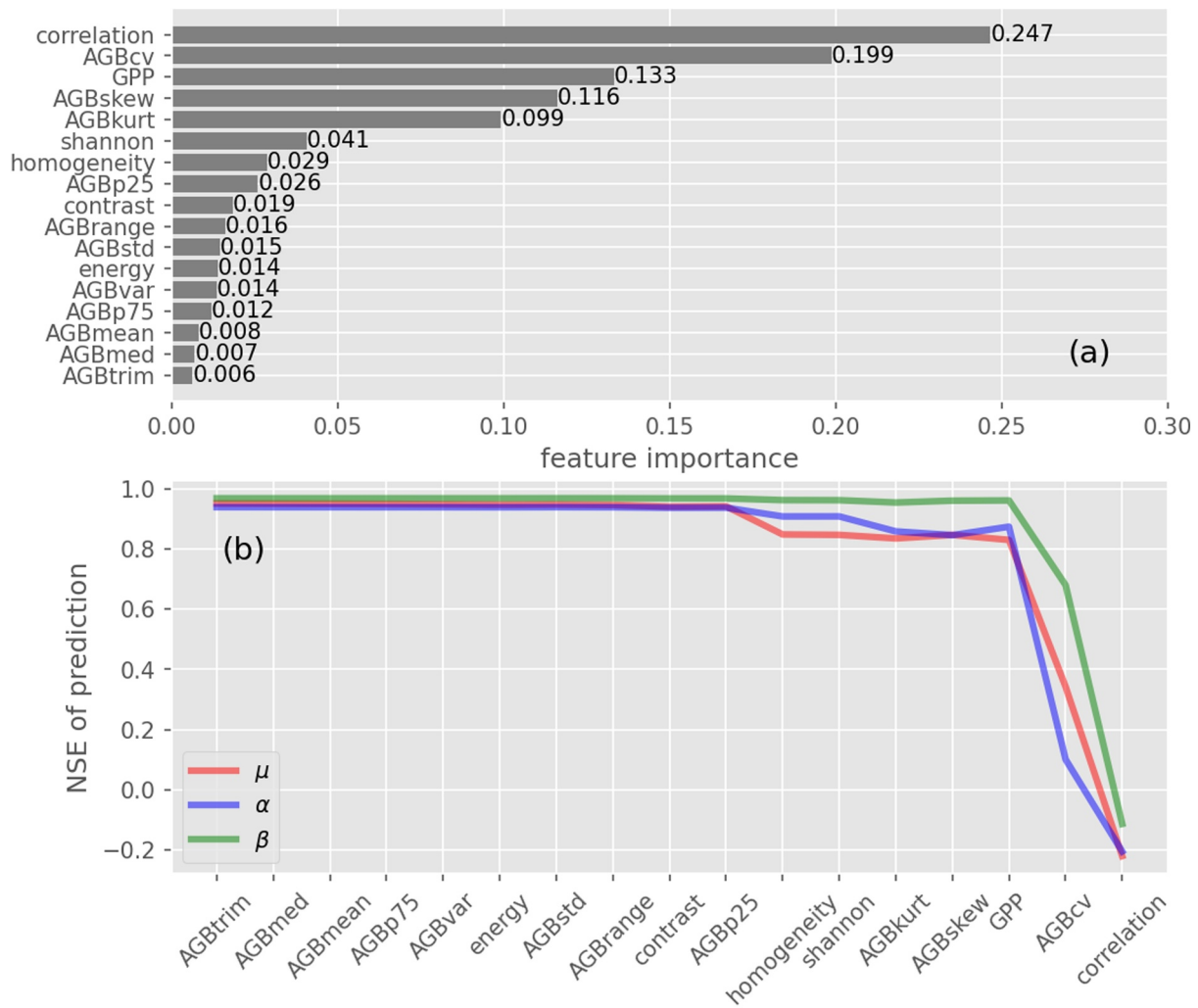
**Figure 6.** Different cross validation accuracy for predicting three disturbance regime parameters. The X-axis denotes the predicted values, and the Y-axis denotes the prescribed values. (a) Illustrates the prediction results of the disturbance regime parameters in the Completely Random cross-validation strategy (CR), (b) refers to the leave one sequence out strategy (LOSO), and (c) refers to the leave one parameter out strategy (LOPO), (d) the LOPO predictions of  $\mu$ ,  $\alpha$ , and  $\beta$  at the boundaries of the prescribed intervals, as well as the values of 0.01 and 0.05 for  $\mu$ , and 0.03 and 0.5 for  $\beta$ .

high degree of precision for estimating  $\mu$  and  $\beta$ , with an NSE approximately 0.82 and 0.85, and a moderately accurate prediction for  $\alpha$  with NSE of 0.69. This suggests that the model has the ability in predicting target disturbance regime parameters when they are not present in the training set, although the level of precision may vary.

The scatter plots of different CV strategies confirmed the high accuracy and apparent gradient for the predictions across the cross-validation strategies (from left to right, Figure 6). The results from the random cross-validation (CR validation, Figure 6a) as well as for the LOPO cross-validation approach (Figure 6b) show robust and high NSE values, and the regression lines close to the 1:1 line, indicating the great accuracy and high correlation with the prescribed values. The LOPO validation (Figure 6b) maintains a similar prediction accuracy with CR validation, but the LOPO results show lower NSE values, larger scatters, and a regression line up to 20% away from the 1:1 line. The performance reduction is especially for parameter  $\alpha$ . One possible reason is that, when training, if certain boundary values of the parameters are absent, the results show an evident bias. To investigate this further, we conduct a test by replacing the LOPO CV predictions by the LOPO CV predictions specially for the  $\mu$ ,  $\alpha$ , and  $\beta$  that fall at the boundaries of the prescribed intervals. As a result, we found a significant increase in precision, with the NSE for  $\mu$  and  $\beta$  exceeding 0.91 and the NSE for  $\alpha$  increasing to 0.76 (Figure 6d). These results confirm the extrapolation challenge faced by our multi-output random forest model when predicting parameter combinations outside the prescribed bounds.

### 3.2.2. Feature Importance

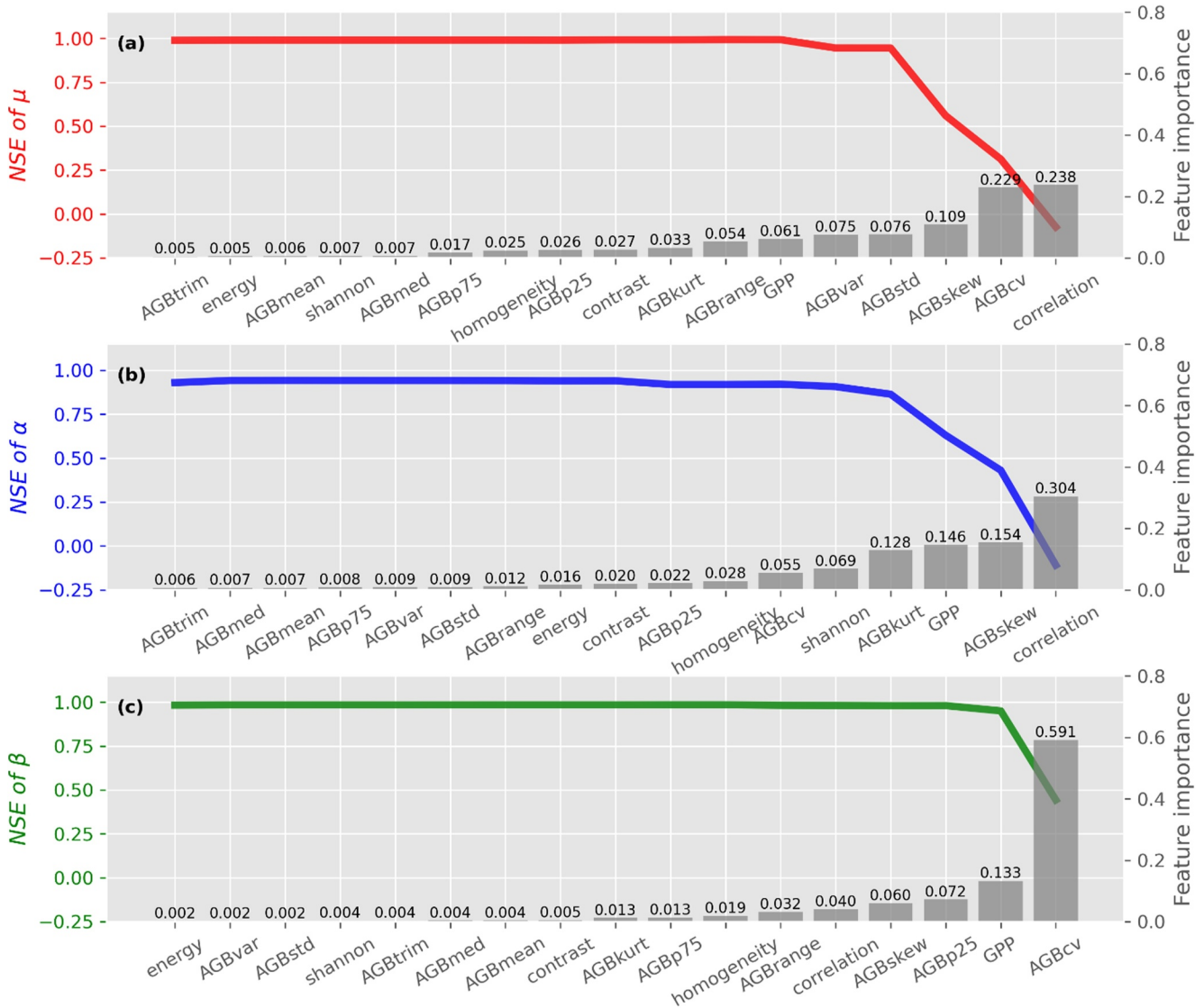
According to the feature importance analysis in the multi-output random forest method, it is observed all types of statistics played significant roles in predicting the disturbance regimes (Figure 7). Additionally, information on



**Figure 7.** (a) Shows the feature importance of multi-output disturbance regime prediction, where the assigned value denotes the degree of contribution made by each feature (see the definition in Table 2). (b) Shows the prediction accuracy change for each disturbance regime parameter, ordered by ascending feature importance. The X-axis represents the feature(s) used (right) and excluded (left) during the prediction process. For instance, when X-axis is labeled as gross primary production (GPP), it means that the prediction process only involved three features, which are GPP, AGBcv, and correlation, and all the other features in the left are excluded. The accuracy was measured using the Nash-Sutcliffe efficiency metric, based on the prescribed parameters and the results of multi-output random forest model's prediction.

mean GPP contributed around 13% for the predictions, ranking it in the third position among the statistical metrics on AGB (Figure 7a). GPP, along with the texture feature correlation, and the coefficient of variation, are the main contributors to the predictions of disturbance regimes (close to 60% of total feature importance). In fact, these three features alone can explain more than 80% of the variation in the different disturbance parameters (Figure 7b).

To elucidate the association between the individual disturbance regime parameters and statistics, we retrained a random forest model specifically for predicting each single parameter. Upon considering features whose contribution exceeds 5%, we see that  $\mu$  is primarily governed by features characterizing the statistical distribution of AGB (feature importance  $\sim$ 55% of the total contribution, Figure 8a). However, texture feature, correlation, is the second most important, accounting for 23% of the contribution, while GPP contributes only 6% for predicting  $\mu$ . The importance of GPP is more apparent in predicting  $\alpha$  (14%) and  $\beta$  (13%). Texture features have a higher significance in predicting parameter  $\alpha$  than  $\mu$  and  $\beta$ , contributing approximately 37%. The feature that made the highest contribution was correlation, which accounted for 30% of the prediction (Figure 8b). For parameter  $\beta$ , the coefficient of variation has a dominant contribution for prediction ( $\sim$ 60%, Figure 8c). The analysis shows that for



**Figure 8.** Breakdown of the feature importance for three individual disturbance regime parameters,  $\mu$ (a),  $\alpha$ (b), and  $\beta$ (c). The corresponding feature importance is depicted through bars, while the colored lines represent the results of the cumulatively exclusive feature test. This test is like Figure 7(b) but employs a single-output random forest model for three parameters individually.

achieving and accuracy over 75% for predicting each of the parameters  $\mu$ ,  $\alpha$ , and  $\beta$  individually, would require the top 4, 3, and 2 features, respectively (Figure 8).

#### 4. Discussion

Disturbance regimes are usually defined by their frequency, severity, and spatial coverage (Liu et al., 2011), and can vary significantly across the landscape (Nelson et al., 1994). Previous approaches have explored the ability to infer different properties of disturbance regimes in several ways using modeling and observations, for example, Fisher et al. (2008) have focused on the number, size, and distribution of disturbance events; Williams et al. (2013) derived disturbance probability and intensity; while Chambers et al. (2013), following Fisher et al. (2008) focused on return frequency, area and tree mortality intensity. A common challenge across the different approaches is the equifinality, the inability to distinguish different disturbance regimes using AGB integrals or only distribution statistics across the landscape or simple regression approaches. Our results show that this challenge can be overcome by exploring higher complexity metrics, using primary productivity for

constraining the problem and exploring machine learning approaches for multi-output regression problems, while some aspects require further discussion.

#### 4.1. Experimental Factorial Design

Building on antecedent research, here we synthesize disturbance regimes in three overarching parameters  $\mu$ ,  $\alpha$ , and  $\beta$ , which control the average area affected by disturbances, their event size-frequency relationship, and the event intensity, respectively. The emergent biomass distribution is further determined by primary productivity, controlled by  $G$ , and by the background mortality,  $K_b$ . We implemented a full factorial design in the variation of these five parameters, imposing a relatively wide range to simulate a sufficiently large sample of disturbance regime scenarios following Fisher et al. (2008) for  $\mu$  and  $\alpha$ , Chambers et al. (2013), for  $\beta$ , Thurner et al. (2016) for  $K_b$ , and allowing  $G$  to generate a range in GPP according to the current ranges emerging from the FLUXNET eddy covariance network (Pastorello et al., 2020).

Although we ensured that the experimental setup spans the parametric ranges found in literature, further findings expanding these intervals should necessarily involve expansion of experimental design. Our study relies on a minimum mortality event size occurring at patch level. The assumption of a patch size of  $10 \times 10 \text{ m}^2$  translates to a maximum domain size of  $10 \times 10 \text{ km}^2$ . Although the patch size, as well as the domain size, can be challenged, namely using Earth observation data, we acknowledge that the current ML model structure is limited to infer disturbances larger than the domain and an occurrence likelihood of one in 200 years. Such aspect is especially relevant given the limitations in predicting boundary values (as shown in Figure 6c). Furthermore, although the event clustering and intensity parameters ( $\alpha$  and  $\beta$ ) span widely, here we follow a prescribed relationship between even size, frequency, and intensity (as in Fisher et al., 2008 and in Chambers et al., 2013). This fixed relationship, despite based on observations (Chambers et al., 2013) could be dependent on disturbance type. Although we include in the current setup an intensity prescription independent of event size ( $\beta = 0.5$ ), different disturbance types could express different average intensities per event and challenge the size-intensity relationship imposed here. The further introduction of disturbance type dependent intensity ranges and relationships brings with it a challenge on equifinality, but should however be considered in future analysis. Alternative formulations on the links between size, frequency and intensity may necessarily lead to different statistical model structures.

##### 4.1.1. Temporal Independence in Disturbance Events

Different types of events may have different size-frequency-intensity relationships, or be temporally correlated (e.g., drought and insect outbreaks (Anderegg et al., 2015)). In this regard, the current prescription of disturbance events is stochastic and, given the lack of quantitative information, temporally independent. The experimental setup is also insensitive to changes in local edaphoclimatic conditions after disturbances. Disturbance events, such as fire, can modify the physical and chemical properties in soils or local microclimatic environments, creating ecological legacies that have cascading implications on carbon dynamics (Liu et al., 2003). The ability to quantify how different disturbances change the posteriori growth conditions and vary the probability of subsequent disturbance events may be very local. However, in a context of climate change and for prognostic purposes, it will support constraining the temporal dynamics of disturbance regimes and potentially provide more realistic projections of carbon cycle responses to climate.

##### 4.1.2. Shapes of Disturbance Events

In our experiments, following also previous studies, for simplicity the disturbance events are prescribed as rectangular shapes (Fisher et al., 2008; Williams et al., 2013). Shapes of disturbance events are usually more complex as demonstrated by high-resolution remote sensing observations (Chambers et al., 2013; Senf & Seidl, 2021b). For most statistical properties, such as distribution or information theory metrics, this aspect is irrelevant. However, the importance of texture features for the prediction of  $\mu$  and  $\alpha$ , especially correlation, may reflect limitations in the generalization of the approach. We additionally conducted a simple exercise including disturbances with irregular shapes to confirm the performance of the approach and that the variable importance is kept (See Figure S3 and Figure S4 of the Supporting Information S1). These results suggest that the landscape texture patterns are mainly controlled by the frequency-size relationship rather than by event shape, and hence type, of disturbances.

#### 4.1.3. Local Biomass Outliers

Occasionally, local and sporadic very high AGB grid cells emerge from the simulations, even after the 200-year simulation period. The outlier, defined as AGB value greater than three standard deviations from the mean value of the column in which the AGB pixel is located (see more details in Text S4 and Table S1 of the Supporting Information S1), was filled with the nearest nonoutlier value. Although increasing the simulation years may theoretically mitigate this phenomenon we found it computationally very inefficient and difficult to control. The presence of such outliers changes some of the statistical features, especially distribution statistics, which leads to a reduction in the ability to predict disturbance regimes. This translates mostly that these local outliers can significantly impact AGB extremes metrics and by that breaking down the relationship between AGB distributions and the underlying disturbance regimes.

#### 4.1.4. Large Scale Controls on Primary Productivity

Another assumption is the dominant role of climate in controlling landscape scale GPP (H. Wang et al., 2014). In our experimental design is the relationship between GPP and biomass, determined by parameter  $G$ . We prescribed a gradient in  $G$  that corresponds to a gradient in maximum GPP at landscape level between 1,000 and 4,000  $\text{gC m}^{-2} \text{yr}^{-1}$ . We acknowledge the heterogeneity of photosynthetic capacity within the domain by randomly perturbing  $\frac{1}{4} G$  for each patch. These perturbations represent a variation of 37% on maximum GPP. The approach falls short by not considering the high frequency environmental controls on photosynthesis, such as solar radiation, temperature, vapor pressure deficit or soil water availability (Bao et al., 2022, 2023). As such, the approach stands on the assumption that landscape scale GPP is mostly controlled by climate rather than by meteorology (Pastorello et al., 2020; H. Wang et al., 2014). Yet, we acknowledge that  $G$  could also be further stratified in space, given the interaction between topography, soils and species composition in providing an additional axis of variation, and consequently warranting further detailed investigation on its role on the link biomass patterns and disturbance regimes. Ultimately, our experimental design translates the strong controls of large scale climate variability on primary productivity and the large mismatch in the spatial resolution of EO-based biomass and photosynthesis data sets, from a few hundred squared meters (e.g.,  $25 \times 25 \text{ m}^2$  GlobBiomass, Santoro et al., 2021) to squared kilometer scales (e.g., MODIS or FLUXCOM GPP products, Jung et al., 2020).

#### 4.2. Performance of the Multi-Output Regression Approach

Overall, the results from the different cross-validation exercises demonstrate the robustness of the approach (Figure 6). The challenges in estimating the disturbance regimes underlying different biomass distributions has been previously highlighted by modeling exercises due to the high spatiotemporal stochasticity that characterizes disturbance events and the equifinality found between disturbance regimes and total biomass distributions (Fisher et al., 2008; Ryan et al., 2012; Williams et al., 2013). We argue that the equifinality issue can be addressed by Equation 1 expanding the feature space to include texture, diversity, and more comprehensive distribution metrics on the emerging biomass patterns; and Equation 2 including additional information about primary productivity. We find that deriving parameters outside the training bounds is a limitation (Figure 6c) but extending the parameter intervals can reduce this problem.

#### 4.3. Landscape Properties Emerging From Disturbances

The multi-output regression shows that the most important variables to predict  $\mu$ ,  $\alpha$ , and  $\beta$  relate to the spatial distribution of biomass, rather than the mean or any higher quantiles of the biomass distribution (Figure 7). Yet, the feature importance rankings change across the three disturbance metrics. The average domain fraction affected by disturbances,  $\mu$ , is strongly linked to biomass distribution statistics (AGBcv, AGBskew, AGBvar, AGBstd) and also to the correlation (Figure 8a). Interestingly biomass itself does not emerge as highly important, although it is implicitly present in AGBcv and AGBstd. The event size-frequency clustering parameter,  $\alpha$ , displays a notable link to text features, such as correlation, homogeneity, and contrast, which account for  $\sim 37\%$  of the overall contributions of all features (Figure 8b). The parameter controlling the relationship between disturbance intensity and event size,  $\beta$ , is mainly dominated by AGBcv (Figure 8c), contributing  $\sim 60\%$  to  $\beta$ 's prediction, followed by GPP ( $\sim 13\%$ ). This connection is the most obvious among the three parameters, indicating that the intensity would directly affect the biomass distribution and indirectly the GPP.

#### 4.4. Linking Modeling Forest Carbon Dynamics With Earth Observation

In this study we used a simple carbon dynamics model to simulate primary productivity and growth resulting from carbon gains and losses at patch level. The realistic annual trajectories in primary productivity and AGB dynamics during recovery (Figure 4), along with model tractability, translate a clear benefit in model simplicity. However, this supports the analysis of the direct disturbances' impact without considering other detailed physiological processes and allocation mechanisms. Realistic biomass dynamics may describe different biomass compartments such as leaves, branches, stems and roots (e.g., CASA, Potter et al., 1993; JSBACH, Reick et al., 2021; DALEC, Williams et al., 2005); differentiate tree density from carbon stocks and include individual and community processes to describe forest dynamics (Bugmann, 2001). Often these models include the effects of climate and other factors on plant growth recovery, such as the impacts of changes in atmospheric CO<sub>2</sub> and in rising temperature (Norby et al., 2001; Pan et al., 2010). The present assumption here is that under given factors promoting primary productivity, disturbances exert the dominant controls on the AGB patterns. Yet we acknowledge the importance of bringing forward more comprehensive mechanistic models with more detailed carbon compartments and plant physiological processes, along with multiple observed constraints for further testing the robustness of these results and for differentiating regional biomass dynamics and corresponding disturbance regimes.

Through the use of remote sensing data and ground-based networks, significant advances have been achieved in understanding, representing, scaling, and characterizing disturbances, ultimately leading to the development of the hypotheses in the process-based models, which can generally be categorized into compartment models and demography models (Liu et al., 2011). The compartment models, including the biogeochemical and ecophysiological ones (Bond-Lamberty et al., 2005; Chen et al., 2000; Liu et al., 2003; McGuire et al., 1992; Parton et al., 1987; Raich et al., 1991; Running & Gower, 1991), can integrate general stand information and meteorological data to simulate carbon cycling, and applied to simulate the biogeochemical processes of forests associated with disturbance (Brugnach, 2005; Tatarinov & Cienciala, 2006; W. Wang et al., 2009). And the demography models, also referred to as gap models, are more focused on the simulation of the impacts of disturbance on the forest composition, structure, and biomass in a relatively long term (Bugmann, 2001; Hurtt et al., 1998; Norby et al., 2001; Shugart et al., 1992). To better capture the fine-tuning functionality of plant sub-compartments and the effects of disturbance on demographic dynamics, the complexity of this type of gap models is increasing (Needham et al., 2022).

Detail in describing disturbances can be also transferred to dynamic global vegetation models (DGVMs), such as LPJ (Haxeltine & Prentice, 1996), HYBRID (Friend et al., 1997), IBIS (Foley et al., 1996), VECODE (Brovkin et al., 1997), and LM3V (Shevliakova et al., 2009). Such could be done, for example, by prescribing only  $\mu$  and intensity losses of carbon at landscape level, or even via lumped parameters (e.g., whole landscape turnover rates), to describe the large-scale impacts of disturbance on carbon dynamics.

Overall, most of these processed-based models at the moment rely on field or satellite observations to quantify and evaluate the impacts of disturbance on modeled carbon stocks or fluxes. But model-data comparisons are far from trivial and often the model-observations mismatch is due to missing information, such as the extent, type, and timing of disturbance events. However, the prescription of individual disturbances based on disturbance regimes metrics derived from observations would minimize this problem and support the analysis of forest dynamics at larger scales. The proliferation of high-resolution biomass from Earth observation, as those by Saatchi et al. (2011) and Santoro et al. (2021), offers a valuable prospect for distinguishing different disturbance regimes.

We would further argue that metrics on disturbance regimes hold information about the different natural and anthropogenic drivers. For example, clearcutting and forest thinning may result in similar spatial patterns of biomass, but with different biomass loss fractions (see Figure S5 a–b of the Supporting Information S1). This could be reflected in the similar  $\mu$  and  $\alpha$ , but different  $\beta$ , which could be the same case for wildfire and drought (see Figure S5 c–d of the Supporting Information S1). However, tornado and insect outbreak (see Figure S5 e–f of the Supporting Information S1) differ in the shape of the affected area, resulting in different biomass clustering patterns, which can be reflected by different  $\alpha$ . Therefore, these natural or anthropogenic factors may cause differential biomass patterns, which are reflected in distinct disturbance parameters.

Furthermore, by analyzing the spatial patterns of biomass stocks over time we can also detect shifts in disturbance regimes and subsequent successional changes. For example, extended dryness periods could induce extensive

drought mortality in some regions, leading to higher values of  $\mu$  and  $\beta$  which would signal transitions in disturbance regimes. Similarly, more frequent wind-throws would result in changes in event clustering patterns over a larger area, presumably indicated by a higher value of  $\mu$  but smaller values of  $\alpha$ . Another example is the intensification in management activities, such as clearcutting, which would cause significant changes in the disturbance regimes, resulting in larger values of  $\mu$ ,  $\alpha$  and  $\beta$ . Information on the transition in disturbance regimes is key for improving our capacity to diagnose relevant changes in forest dynamics and for understanding the relationship between vegetation mortality, the carbon cycle dynamics and climate variability. It can assist us in identifying susceptible regions that are particularly vulnerable to disturbances, thereby facilitating the implementation of effective climate adaptation and mitigation strategies, but also management strategies and conservation efforts to mitigate the impacts of disturbance on both ecosystem functioning and biodiversity.

## 5. Conclusion

This study presents a framework for simulating disturbance events using a wide range of disturbance regime attributes. Together with a simple carbon dynamics model, we simulated the effects of disturbance events, from combinations of three landscape-level disturbance regime parameters:  $\mu$  (probability scale),  $\alpha$  (clustering degree), and  $\beta$  (intensity slope) on biomass patterns. We observe how changes in extent, frequency, and intensity characterizing the different disturbance regimes indeed lead to significantly different biomass patterns even for analogous primary productivity and background mortality inputs.

The emerging spatial biomass patterns are summarized in a set of metrics containing the central tendencies, moments, as well as information-based and texture information. Based on a conceptual model-based experiment and machine learning regression we demonstrate that with this set of summary metrics of biomass, along with primary productivity constraints, the set of disturbance regime parameters can be reliably retrieved. The average fraction of the domain affected by disturbances, the event size clustering exponent, and the perturbation intensity could be retrieved with an accuracy of 94.8%, 94.9%, and 97.1%, respectively.

As Earth observation efforts evolve to retrieve space-borne estimates of vegetation structure and patterns such as GEDI (Stavros et al., 2017), NISAR (Rosen et al., 2015), and BIOMASS (Le Toan et al., 2011), and as photosynthesis patterns are being estimated at high resolution (Cogliati et al., 2015; Jung et al., 2020), the methods presented in this study will open up avenues to provide observation based long-term diagnostics on the terrestrial carbon cycle and background disturbance patterns, which could be used to constrain Earth system models.

## Conflict of Interest

The authors declare that they have no conflict of interest.

## Data Availability Statement

The framework of generating reference disturbance data set was developed in MATLAB 2022a, available under MATLAB license at <https://www.mathworks.com/products/matlab.html>. The multi-output machine learning regression approach and the visualizations were implemented in Python 3.8. All scripts executed to yield the results of this article are available at Wang (2023).

## References

- Allen, C., Macalady, A., Bachelet, D., McDowell, N., Venetier, M., Kitzberger, T., et al. (2010). A global overview of drought and heat-induced tree mortality reveals emerging climate change risks for forests. *Forest Ecology and Management*, 259(4), 660–684. <https://doi.org/10.1016/j.foreco.2009.09.001>
- Amthor, J. S. (2000). The McCree–de Wit–Penning de Vries–Thornley Respiration Paradigms: 30 Years Later. *Annals of Botany*, 86(1), 1–20. <https://doi.org/10.1006/anbo.2000.1175>
- Anderegg, W. R. L., Hicke, J. A., Fisher, R. A., Allen, C. D., Aukema, J., Bentz, B., et al. (2015). Tree mortality from drought, insects, and their interactions in a changing climate. *New Phytologist*, 208(3), 674–683. <https://doi.org/10.1111/nph.13477>
- Araujo, R. F., Grubinger, S., Celes, C. H. S., Negrón-Juárez, R. I., Garcia, M., Dandois, J. P., & Muller-Landau, H. C. (2021). Strong temporal variation in treefall and branchfall rates in a tropical forest is related to extreme rainfall: Results from 5 years of monthly drone data for a 50 ha plot. *Biogeosciences*, 18(24), 6517–6531. <https://doi.org/10.5194/bg-18-6517-2021>
- Bao, S., Alonso, L., Wang, S., Gensheimer, J., De, R., & Carvalhais, N. (2023). Toward robust parameterizations in ecosystem-level photosynthesis models. *Journal of Advances in Modeling Earth Systems*, 15(8), e2022MS003464. <https://doi.org/10.1029/2022MS003464>
- Bao, S., Wutzler, T., Koirala, S., Cuntz, M., Ibrom, A., Besnard, S., et al. (2022). Environment-sensitivity functions for gross primary productivity in light use efficiency models. *Agricultural and Forest Meteorology*, 312, 108708. <https://doi.org/10.1016/j.agrformet.2021.108708>

## Acknowledgments

The authors express their gratitude to the Associate Editor and two anonymous Reviewers for their constructive feedback on a first version of the manuscript. SW acknowledges support from the International Max Planck Research School for Biogeochemical Cycles (IMPRS-gBGC). SW and HY acknowledge EEBIOMASS, Project Office BIOMASS, support from the German Federal Ministry of Economics and Technology (50EE1904). Open Access funding enabled and organized by Projekt DEAL.



- Bond-Lamberty, B., Gower, S. T., Ahl, D. E., & Thornton, P. E. (2005). Reimplementation of the Biome-BGC model to simulate successional change. *Tree Physiology*, 25(4), 413–424. <https://doi.org/10.1093/treephys/25.4.413>
- Bossel, H., & Krieger, H. (1994). Simulation of multi-species tropical forest dynamics using a vertically and horizontally structured model. *Forest Ecology and Management*, 69(1), 123–144. [https://doi.org/10.1016/0378-1127\(94\)90224-0](https://doi.org/10.1016/0378-1127(94)90224-0)
- Brokaw, N. V. L., & Scheiner, S. M. (1989). Species composition in gaps and structure of a tropical forest. *Ecology*, 70(3), 538–541. <https://doi.org/10.2307/1940196>
- Brovkin, V., Ganopol'ski, A., & Svirezhev, Y. (1997). A continuous climate-vegetation classification for use in climate-biosphere studies. *Ecological Modelling*, 101(2), 251–261. [https://doi.org/10.1016/S0304-3800\(97\)00049-5](https://doi.org/10.1016/S0304-3800(97)00049-5)
- Brugnach, M. (2005). Process level sensitivity analysis for complex ecological models. *Ecological Modelling*, 187(2), 99–120. <https://doi.org/10.1016/j.ecolmodel.2005.01.044>
- Bugmann, H. (2001). A review of forest gap models. *Climatic Change*, 51(3), 259–305. <https://doi.org/10.1023/A:1012525626267>
- Carvalho, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., et al. (2014). Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature*, 514(7521), 213–217. <https://doi.org/10.1038/nature13731>
- Chambers, J. Q., Higuchi, N., Teixeira, L. M., dos Santos, J., Laurance, S. G., & Trumbore, S. E. (2004). Response of tree biomass and wood litter to disturbance in a Central Amazon forest. *Oecologia*, 141(4), 596–611. <https://doi.org/10.1007/s00442-004-1676-2>
- Chambers, J. Q., Negron-Juarez, R. I., Marra, D. M., Vittorio, A. D., Tews, J., Roberts, D., et al. (2013). The steady-state mosaic of disturbance and succession across an old-growth Central Amazon forest landscape. *Proceedings of the National Academy of Sciences*, 110(10), 3949–3954. <https://doi.org/10.1073/pnas.1202894110>
- Chen, W., Chen, J., & Cihlar, J. (2000). An integrated terrestrial ecosystem carbon-budget model based on changes in disturbance, climate, and atmospheric chemistry. *Ecological Modelling*, 135(1), 55–79. [https://doi.org/10.1016/S0304-3800\(00\)00371-9](https://doi.org/10.1016/S0304-3800(00)00371-9)
- Cogliati, S., Verhoef, W., Kraft, S., Sabater, N., Alonso, L., Vicent, J., et al. (2015). Retrieval of sun-induced fluorescence using advanced spectral fitting methods. *Remote Sensing of Environment*, 169, 344–357. <https://doi.org/10.1016/j.rse.2015.08.022>
- Delbart, N., Ciais, P., Chave, J., Viovy, N., Malhi, Y., & Le Toan, T. (2010). Mortality as a key driver of the spatial distribution of aboveground biomass in Amazonian forest: Results from a dynamic vegetation model. *Biogeosciences*, 7(10), 3027–3039. <https://doi.org/10.5194/bg-7-3027-2010>
- Fisher, J. I., Hurr, G. C., Thomas, R. Q., & Chambers, J. Q. (2008). Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecology Letters*, 11(6), 554–563. <https://doi.org/10.1111/j.1461-0248.2008.01169.x>
- Foley, J. A., Prentice, I. C., Ramankutty, N., Levis, S., Pollard, D., Sitch, S., & Haxeltine, A. (1996). An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles*, 10(4), 603–628. <https://doi.org/10.1029/96GB02692>
- Franklin, J. F., Shugart, H. H., & Harmon, M. E. (1987). Tree death as an ecological process. *BioScience*, 37(8), 550–556. <https://doi.org/10.2307/1310665>
- Friend, A. D., Lucht, W., Rademacher, T. T., Kerbin, R., Betts, R., Cadule, P., et al. (2014). Carbon residence time dominates uncertainty in terrestrial vegetation responses to future climate and atmospheric CO<sub>2</sub>. *Proceedings of the National Academy of Sciences*, 111(9), 3280–3285. <https://doi.org/10.1073/pnas.1222477110>
- Friend, A. D., Stevens, A. K., Knox, R. G., & Cannell, M. G. R. (1997). A process-based, terrestrial biosphere model of ecosystem dynamics (Hybrid v3.0). *Ecological Modelling*, 95(2), 249–287. [https://doi.org/10.1016/S0304-3800\(96\)00034-8](https://doi.org/10.1016/S0304-3800(96)00034-8)
- Grime, J. P. (1977). Evidence for the existence of three primary strategies in plants and its relevance to ecological and evolutionary theory. *The American Naturalist*, 111(982), 1169–1194. <https://doi.org/10.1086/283244>
- Hammond, W. M., Williams, A. P., Abatzoglou, J. T., Adams, H. D., Klein, T., López, R., et al. (2022). Global field observations of tree die-off reveal hotter-drought fingerprint for Earth's forests. *Nature Communications*, 13(1), 1761. Article 1. <https://doi.org/10.1038/s41467-022-29289-2>
- Haralick, R. M., Shanmugam, K., & Dinstein, I. (1973). Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6), 610–621. <https://doi.org/10.1109/TSMC.1973.4309314>
- Haxeltine, A., & Prentice, I. C. (1996). BIOME3: An equilibrium terrestrial biosphere model based on ecophysiological constraints, resource availability, and competition among plant functional types. *Global Biogeochemical Cycles*, 10(4), 693–709. <https://doi.org/10.1029/96GB02344>
- Hill, T., Ryan, C., & Williams, M. (2015). A framework for estimating forest disturbance intensity from successive remotely sensed biomass maps: Moving beyond average biomass loss estimates. *Carbon Balance and Management*, 10(1), 27. <https://doi.org/10.1186/s13021-015-0039-0>
- Hurr, G. C., Moorcroft, P. L. R., And, S. W. P., & Levin, S. A. (1998). Terrestrial models and global change: Challenges for the future. *Global Change Biology*, 4(5), 581–590. <https://doi.org/10.1046/j.1365-2486.1998.t01-1-00203.x>
- Jans, L., Poorter, L., van Rompaey, R. S. A. R., & Bongers, F. (1993). Gaps and forest zones in tropical moist forest in ivory coast. *Biotropica*, 25(3), 258–269. <https://doi.org/10.2307/2388784>
- Johnson, M. O., Galbraith, D., Gloor, M., De Duerwaerder, H., Guimberteau, M., Rammig, A., et al. (2016). Variation in stem mortality rates determines patterns of above-ground biomass in Amazonian forests: Implications for dynamic global vegetation models. *Global Change Biology*, 22(12), 3996–4013. <https://doi.org/10.1111/gcb.13315>
- Jung, M., Schwalm, C., Migliavacca, M., Walther, S., Camps-Valls, G., Koirala, S., et al. (2020). Scaling carbon fluxes from eddy covariance sites to globe: Synthesis and evaluation of the FLUXCOM approach. *Biogeosciences*, 17(5), 1343–1365. <https://doi.org/10.5194/bg-17-1343-2020>
- Köhler, P., & Huth, A. (1998). The effects of tree species grouping in tropical rainforest modelling: Simulations with the individual-based model Formind. *Ecological Modelling*, 109(3), 301–321. [https://doi.org/10.1016/S0304-3800\(98\)00066-0](https://doi.org/10.1016/S0304-3800(98)00066-0)
- Lawton, R. O., & Putz, F. E. (1988). Natural disturbance and gap-phase regeneration in a wind-exposed tropical cloud forest. *Ecology*, 69(3), 764–777. <https://doi.org/10.2307/1941025>
- Le Toan, T., Quegan, S., Davidson, M. W. J., Baltzer, H., Paillou, P., Papatianassiou, K., et al. (2011). The BIOMASS mission: Mapping global forest biomass to better understand the terrestrial carbon cycle. *Remote Sensing of Environment*, 115(11), 2850–2860. <https://doi.org/10.1016/j.rse.2011.03.020>
- Liu, S., Bliss, N., Sundquist, E., & Huntington, T. G. (2003). Modeling carbon dynamics in vegetation and soil under the impact of soil erosion and deposition. *Global Biogeochemical Cycles*, 17(2). <https://doi.org/10.1029/2002GB002010>

- Liu, S., Bond-Lamberty, B., Hicke, J. A., Vargas, R., Zhao, S., Chen, J., et al. (2011). Simulating the impacts of disturbances on forest carbon cycling in North America: Processes, data, models, and challenges. *Journal of Geophysical Research*, *116*(G4), G00K08. <https://doi.org/10.1029/2010JG001585>
- Malhi, Y., Phillips, O. L., Phillips, O. L., Baker, T. R., Arroyo, L., Higuchi, N., et al. (2004). Pattern and process in Amazon tree turnover, 1976–2001. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, *359*(1443), 381–407. <https://doi.org/10.1098/rstb.2003.1438>
- McDowell, N. G., Sapes, G., Pivovarov, A., Adams, H. D., Allen, C. D., Anderegg, W. R. L., et al. (2022). Mechanisms of woody-plant mortality under rising drought, CO<sub>2</sub> and vapour pressure deficit. *Nature Reviews Earth and Environment*, *3*(5), 1–15. <https://doi.org/10.1038/s43017-022-00272-1>
- McGuire, A. D., Melillo, J. M., Joyce, L. A., Kicklighter, D. W., Grace, A. L., Moore III, B., & Vorosmarty, C. J. (1992). Interactions between carbon and nitrogen dynamics in estimating net primary productivity for potential vegetation in North America. *Global Biogeochemical Cycles*, *6*(2), 101–124. <https://doi.org/10.1029/92GB00219>
- Moorecroft, P. R., Hurtt, G. C., & Pacala, S. W. (2001). A method for scaling vegetation dynamics: The ecosystem demography model (ED). *Ecological Monographs*, *71*(4), 557–586. [https://doi.org/10.1890/0012-9615\(2001\)071\[0557:AMFSVD\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2001)071[0557:AMFSVD]2.0.CO;2)
- Moriyas, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900. <https://doi.org/10.13031/2013.23153>
- Nash, J. E., & Sutcliffe, J. V. (1970). River flow forecasting through conceptual models part I—A discussion of principles. *Journal of Hydrology*, *10*(3), 282–290. [https://doi.org/10.1016/0022-1694\(70\)90255-6](https://doi.org/10.1016/0022-1694(70)90255-6)
- Needham, J. F., Arellano, G., Davies, S. J., Fisher, R. A., Hammer, V., Knox, R. G., et al. (2022). Tree crown damage and its effects on forest carbon cycling in a tropical forest. *Global Change Biology*, *28*(18), 5560–5574. <https://doi.org/10.1111/gcb.16318>
- Nelson, B. W., Kapos, V., Adams, J. B., Oliveira, W. J., & Braun, O. P. G. (1994). Forest disturbance by large blowdowns in the Brazilian Amazon. *Ecology*, *75*(3), 853–858. <https://doi.org/10.2307/1941742>
- Norby, R. J., Ogle, K., Curtis, P. S., Badeck, F.-W., Huth, A., Hurtt, G. C., et al. (2001). Aboveground growth and competition in forest gap models: An analysis for studies of climatic change. *Climatic Change*, *51*(3), 415–447. <https://doi.org/10.1023/A:1012510619424>
- Pan, Y., Chen, J. M., Birdsey, R., McCullough, K., He, L., & Deng, F. (2010). Age structure and disturbance legacy of North American forests. *Biogeosciences Discussions*, *7*, 979–1020. <https://doi.org/10.5194/bgd-7-979-2010>
- Parton, W. J., Schimel, D. S., Cole, C. V., & Ojima, D. S. (1987). Analysis of factors controlling soil organic matter levels in great plains grasslands. *Soil Science Society of America Journal*, *51*(5), 1173–1179. <https://doi.org/10.2136/sssaj1987.03615995005100050015x>
- Pastorello, G., Trotta, C., Canfora, E., Chu, H., Christianson, D., Cheah, Y.-W., et al. (2020). The FLUXNET2015 dataset and the ONEFlux processing pipeline for eddy covariance data. *Scientific Data*, *7*(1), 225. Article 1. <https://doi.org/10.1038/s41597-020-0534-3>
- Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, *12*(85), 2825–2830.
- Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*(4), 811–841. <https://doi.org/10.1029/93GB02725>
- Raich, J. W., Rastetter, E. B., Melillo, J. M., Kicklighter, D. W., Steudler, P. A., Peterson, B. J., et al. (1991). Potential net primary productivity in south America: Application of a global model. *Ecological Applications*, *1*(4), 399–429. <https://doi.org/10.2307/1941899>
- Reick, C. H., Gayler, V., Goll, D., Hagemann, S., Heidkamp, M., Nabel, J. E. M. S., et al. (2021). JSBACH 3 - the land component of the MPI Earth system model: Documentation of version 3.2. <https://doi.org/10.17617/2.3279802>
- Rosen, P. A., Hensley, S., Shaffer, S., Veilleux, L., Chakraborty, M., Misra, T., et al. (2015). The NASA-ISRO SAR mission—An international space partnership for science and societal benefit. In *2015 IEEE radar conference (RadarCon)* (pp. 1610–1613). IEEE. <https://doi.org/10.1109/RADAR.2015.7131255>
- Runkle, J. R. (2000). Canopy tree turnover in old-growth mesic forests of eastern north America. *Ecology*, *81*(2), 554–567. <https://doi.org/10.2307/177448>
- Running, S. W., & Gower, S. T. (1991). FOREST-BGC, A general model of forest ecosystem processes for regional applications. II. Dynamic carbon allocation and nitrogen budgets I. *Tree Physiology*, *9*(1–2), 147–160. <https://doi.org/10.1093/treephys/9.1-2.147>
- Ryan, C. M., Hill, T., Woollen, E., Ghee, C., Mitchard, E., Cassells, G., et al. (2012). Quantifying small-scale deforestation and forest degradation in African woodlands using radar imagery. *Global Change Biology*, *18*(1), 243–257. <https://doi.org/10.1111/j.1365-2486.2011.02551.x>
- Ryan, C. M., & Williams, M. (2011). How does fire intensity and frequency affect miombo woodland tree populations and biomass? *Ecological Applications*, *21*(1), 48–60. <https://doi.org/10.1890/09-1489.1>
- Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T. A., Salas, W., et al. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences of the United States of America*, *108*(24), 9899–9904. <https://doi.org/10.1073/pnas.1019576108>
- Santoro, M., Cartus, O., Carvalhais, N., Rozendaal, D. M. A., Avitabile, V., Araza, A., et al. (2021). The global forest above-ground biomass pool for 2010 estimated from high-resolution satellite observations. *Earth System Science Data*, *13*(8), 3927–3950. <https://doi.org/10.5194/essd-13-3927-2021>
- Santoro, M., Cartus, O., Wegmüller, U., Besnard, S., Carvalhais, N., Araza, A., et al. (2022). Global estimation of above-ground biomass from spaceborne C-band scatterometer observations aided by LiDAR metrics of vegetation structure. *Remote Sensing of Environment*, *279*, 113114. <https://doi.org/10.1016/j.rse.2022.113114>
- Senf, C., & Seidl, R. (2021a). Mapping the forest disturbance regimes of Europe. *Nature Sustainability*, *4*(1), 63–70. <https://doi.org/10.1038/s41893-020-00609-y>
- Senf, C., & Seidl, R. (2021b). Storm and fire disturbances in Europe: Distribution and trends. *Global Change Biology*, *27*(15), 3605–3619. <https://doi.org/10.1111/gcb.15679>
- Shannon, C. E. (1948). A mathematical theory of communication. *Bell System Technical Journal*, *27*(3), 379–423. <https://doi.org/10.1002/j.1538-7305.1948.tb01338.x>
- Shevliakova, E., Pacala, S. W., Malyshev, S., Hurtt, G. C., Milly, P. C. D., Caspersen, J. P., et al. (2009). Carbon cycling under 300 years of land use change: Importance of the secondary vegetation sink. *Global Biogeochemical Cycles*, *23*(2). <https://doi.org/10.1029/2007GB003176>
- Shugart, H. H., Leemans, R., & Bonan, G. B. (Eds.) (1992). *A systems analysis of the global boreal forest*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511565489>
- Spellerberg, I. F., & Fedor, P. J. (2003). A tribute to Claude Shannon (1916–2001) and a plea for more rigorous use of species richness, species diversity and the ‘Shannon–Wiener’ Index. *Global Ecology and Biogeography*, *12*(3), 177–179. <https://doi.org/10.1046/j.1466-822X.2003.00015.x>

- Stavros, E. N., Schimel, D., Pavlick, R., Serbin, S., Swann, A., Duncanson, L., et al. (2017). ISS observations offer insights into plant function. *Nature Ecology & Evolution*, *1*(7), 0194. Article 7. <https://doi.org/10.1038/s41559-017-0194>
- Tatarinov, F. A., & Cienciala, E. (2006). Application of BIOME-BGC model to managed forests: 1. Sensitivity analysis. *Forest Ecology and Management*, *237*(1), 267–279. <https://doi.org/10.1016/j.foreco.2006.09.085>
- Thurner, M., Beer, C., Carvalhais, N., Forkel, M., Santoro, M., Tum, M., & Schmillius, C. (2016). Large-scale variation in boreal and temperate forest carbon turnover rate related to climate. *Geophysical Research Letters*, *43*(9), 4576–4585. <https://doi.org/10.1002/2016GL068794>
- Turner, M. G. (2010). Disturbance and landscape dynamics in a changing world. *Ecology*, *91*(10), 2833–2849. <https://doi.org/10.1890/10-0097.1>
- Wang, H., Prentice, I. C., & Davis, T. W. (2014). Biophysical constraints on gross primary production by the terrestrial biosphere. *Biogeosciences*, *11*(20), 5987–6001. <https://doi.org/10.5194/bg-11-5987-2014>
- Wang, S. (2023). RS-SiyuanWang/Disturbance-Regime: Updates for peer-reviewing (v1.2) [Dataset]. *Zenodo*. <https://doi.org/10.5281/zenodo.8121119>
- Wang, W., Ichii, K., Hashimoto, H., Michaelis, A. R., Thornton, P. E., Law, B. E., & Nemani, R. R. (2009). A hierarchical analysis of terrestrial ecosystem model Biome-BGC: Equilibrium analysis and model calibration. *Ecological Modelling*, *220*(17), 2009–2023. <https://doi.org/10.1016/j.ecolmodel.2009.04.051>
- Williams, M., Hill, T. C., & Ryan, C. M. (2013). Using biomass distributions to determine probability and intensity of tropical forest disturbance. *Plant Ecology and Diversity*, *6*(1), 87–99. <https://doi.org/10.1080/17550874.2012.692404>
- Williams, M., Schwarz, P. A., Law, B. E., Irvine, J., & Kurpius, M. R. (2005). An improved analysis of forest carbon dynamics using data assimilation. *Global Change Biology*, *11*(1), 89–105. <https://doi.org/10.1111/j.1365-2486.2004.00891.x>
- Yan, Z., Li, J., Fei, J., Mao, X., Gao, P., & Ding, Y. (2005). Study on the adsorptive catalytic voltammetry of emodin at a carbon paste electrode. *Analytical Letters*, *38*(10), 1641–1650. <https://doi.org/10.1081/AL-200065808>
- Yavitt, J., Battles, J., Lang, G. E., & Knight, D. H. (1995). The canopy gap regime in a secondary Neotropical forest in Panama. *Journal of Tropical Ecology*, *11*(3), 391–402. <https://doi.org/10.1017/S0266467400008853>