Understanding disturbance regimes from patterns in biomass and primary productivity

Siyuan Wang[1,1], Hui Yang[1,1], Sujan Koirala[1,1], Matthias Forkel[2,2], Markus Reichstein[1,1], and Nuno Carvalhais[1,1]

[1]Max-Planck Institute for Biogeochemistry
[2]TU Dresden

November 30, 2022

Abstract

Natural and anthropogenic disturbances act as important drivers of tree mortality, shaping the structure, composition and biomass distribution of forests. Disturbance regimes may emerge from different characteristics of disturbance events over time and space. We design a model- based experiment to investigate the links between disturbance regimes at the landscape scale and spatial features of biomass patterns. The effects on biomass of a wide range of disturbance regimes are simulated by varying three different parameters, i.e. μ (probability scale), α (clustering degree), and β (intensity slope) that shape the extent, frequency, and intensity of disturbance events, respectively. A simple dynamic carbon cycle model is used to simulate 200 years of plant biomass dynamics in response to circa +2000 different disturbance regimes, depending on the different combinations of μ, α, and β. Each parameter combination yields a spatially explicit estimate of plant biomass for which sixteen synthesis statistics are estimated on the spatial distributions of biomass, including information-based and texture features. Based on a multi-output regression approach we link these synthesis statistics with additional gross primary production (GPP) constraints to retrieve the three disturbance parameters. In doing so we evaluate the confidence in inferring disturbance regimes from spatial distributions of biomass. Our results show that all three parameters can be confidently retrieved. The Nash-Sutcliffe efficiency for the prediction of the μ, α, and β is 97.3%, 96.6%, and 97.9%, respectively. A feature importance analysis reveals that the distribution statistics dominate the prediction of μ and β, while features quantifying texture have a stronger connection with α. Overall, this study clarifies the association between biomass patterns emerging from different underlying disturbance regimes, while overcoming the previously found equifinality between mortality rates and total biomass. Given the links between decadal vegetation dynamics and the uncertainties in the role of terrestrial ecosystems in the global biogeochemical cycles, a better understanding and the quantification of disturbance regimes would improve our current understanding of controls and feedback at the biosphere-atmosphere interface in the current Earth system models.

Understanding disturbance regimes from patterns in biomass and primary productivity

Siyuan Wang[1,2], Hui Yang[1], Sujan Koirala[1], Matthias Forkel[2], Markus Reichstein[1,3], Nuno Carvalhais[1,3,4*]*

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

(*Corresponding authors: Siyuan Wang, swang@bgc-jena.mpg.de, Nuno Carvalhais, ncarvalhais@bgc-jena.mpg.de)

Abstract

Natural and anthropogenic disturbances act as important drivers of tree mortality, shaping the structure, composition and biomass distribution of forests. Disturbance regimes may emerge from different characteristics of disturbance events over time and space. We design a model-based experiment to investigate the links between disturbance regimes at the landscape scale and spatial features of biomass patterns. The effects on biomass of a wide range of disturbance regimes are simulated by varying three different parameters, i.e. μ (probability scale), α (clustering degree), and β (intensity slope) that shape the extent, frequency, and intensity of disturbance events, respectively. A simple dynamic carbon cycle model is used to simulate 200 years of plant biomass dynamics in response to circa +2000 different disturbance regimes, depending on the different combinations of μ, α, and β. Each parameter combination yields a spatially explicit estimate of plant biomass for which sixteen synthesis statistics are estimated on the spatial distributions of biomass, including information-based and texture features. Based on a multi-output regression approach we link these synthesis statistics with additional gross primary production (GPP) constraints to retrieve the three disturbance parameters. In doing so we evaluate the confidence in inferring disturbance regimes from spatial distributions of biomass. Our results show that all three parameters can be confidently retrieved. The Nash-Sutcliffe efficiency for the prediction of the μ, α, and β is 97.3%, 96.6%, and 97.9%, respectively. A feature importance analysis reveals that the distribution statistics dominate the prediction of μ and β, while features quantifying texture have a stronger connection with α. Overall, this study clarifies the association between biomass patterns emerging from different underlying disturbance regimes, while overcoming the previously found equifinality between mortality rates and total biomass. Given the links between decadal vegetation dynamics and the uncertainties in the role of terrestrial ecosystems in the global biogeochemical cycles, a better understanding and the quantification of disturbance regimes would improve our current understanding of controls and feedback at the biosphere-atmosphere interface in the current Earth system models.

1 Introduction

Mortality is one of the key processes in vegetation dynamics (Franklin, Shugart, and Harmon 1987; Runkle 2000), which dominates the aboveground carbon turnover rate (Carvalhais et al. 2014; Thurner et al. 2016) and contributes to significant uncertainties in carbon fluxes and carbon budget (Andrew D. Friend et al. 2014). The diverse range of natural (e.g. fire, droughts, windthrows, pathogens, and insects) and anthropogenic disturbances (e.g. agriculture, urbanization, and clearcutting) act as strong drivers of vegetation mortality, leading to the total or partial loss of biomass (McDowell et al. 2022; Hammond et al. 2022; Grime 1977). A better understanding of the mortality caused by disturbance, as well as its impacts on vegetation carbon dynamics, is crucial for constraining future carbon cycling prognostics (Andrew D. Friend et al. 2014).

Comparatively to primary productivity and allocation, mortality caused by disturbances also plays an equivalently important role in determining the spatial gradients of above-ground biomass (AGB, Delbart et al. 2010) and is currently poorly monitored due to its highly stochastic nature (Chambers et al. 2004; Allen et al. 2010; Hammond et al. 2022). Characteristics of these stochastic disturbance events at the landscape level, also called disturbance regimes, are commonly interpreted as different metrics, such as size, frequency, intensity, and aggregation (Turner 2010). These different attributes of disturbance regimes are the description of cumulative effects of all disturbance events occurring in a given area and time period (Senf and Seidl 2021b), ultimately leading to a shifting steady-state mosaic, represented by patches of distinct successional ages or carbon stocks over long time periods (Brokaw and Scheiner 1989).

Most research on quantifying disturbance regime parameters has been carried out either in an observation-driven manner or using model-data-integration simulation. Through large scale time-series remote sensing data, the information on vegetation change, such as spectral bands/indices and segment outline is able to be directly extracted for describing the disturbance magnitude, duration, and rate of change (Chambers et al. 2013; Senf and Seidl 2021b). On the other hand, the disturbance regime can also be simulated beforehand and incorporated into the process- and individual-based models (Bugmann 2001; Bossel and Krieger 1994; Yan et al. 2005; Köhler and Huth 1998), in which the relationship between disturbance regime parameters and model outputs, such as above-ground biomass, can be expected to be established. Previous research has revealed that two disturbance parameters that determine the average probability and intensity of biomass loss, can be retrieved back via satellite biomass observation (Williams, Hill, and Ryan 2013). Further relying on successive biomass maps, the difference in patterns of intensity ranging from deforestation to widespread low-intensity disturbance can also be detected (Hill, Ryan, and Williams 2015), which indicate a strong relationship between disturbance regimes and biomass observations.

It should be noted that the clustering pattern of disturbance events has been recognized as a fundamental attribute of the disturbance regime. And if these disturbance events are not correctly characterized, the average mortality and growth would lead to misestimations (Fisher et al. 2008). Therefore, caution is advised for considering the different disturbance events

clustering patterns in the process of simulation. Besides, most previous works about simulating disturbance events were under relatively similar growth processes, which simplified the heterogeneity of the forest's primary productivity. The development of high-resolution biomass observations (Saatchi et al. 2011; Santoro et al. 2021), indicates a need to understand more comprehensive disturbance regimes through biomass maps.

In this paper, we design a model-based experiment to investigate the links between disturbance regimes at the landscape scale and spatial features of biomass patterns. The effects on biomass of a wide range of disturbance regimes are simulated based on different μ (probability scale), α (clustering degree), and β (intensity slope) that respectively shape the extent, frequency, and intensity of disturbance events. A simple dynamic carbon cycle model is used to simulate 200 years of plant biomass dynamics in response to circa +2000 different disturbance regimes, depending on different combinations of μ, α, and β. Each parameter combination yields a spatially explicit estimate of plant biomass for which different synthesis statistics are estimated (e.g., mean, median, standard deviation, quantiles, and skewness). Based on a multi-output regression approach we link these synthesis statistics back to the three disturbance parameters to evaluate the confidence in inferring disturbance regimes from spatial distributions of biomass alone.

2 Method

2.1 Dynamic carbon cycle model

We determined the dynamic processes of carbon cycling at the patch level. Each patch is specified as a square pixel with a length of 10m, which means the smallest computing unit corresponds to a 0.01ha plant crown area (Fisher et al. 2008). The changes of vegetation carbon (C, in $kgC \cdot m^{-2} \cdot yr^{-1}$) are controlled by the differences between carbon gains (via photosynthesis, NPP_{AGB}) and losses (L, Eq. 1). Here the simulations are performed at annual timescales, which allow rapid computations and flexibility on the occurrence of disturbance events.

$$\frac{\mathrm{d}AGB}{\mathrm{d}t} = NPP_{AGB} - L \qquad \textit{Eq.1}$$

For carbon gain part (NPP_{AGB}), it comes from the gross primary production (GPP) which loses a fraction of C to growth respiration ($1 - Y_G$, Amthor 2000), and part of which is allocated to above-ground biomass (f_{AGB}) as follow:

$$NPP_{AGB} = GPP \times (1 - Y_G) \times f_{AGB} \qquad \textit{Eq.2}$$

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 allocation of C to the above ground pools, and a growth respiration ratio of 0.25 (Amthor 2000). And we performed a simple relationship between GPP and biomass, which can be mathematically described as

$$GPP = \frac{100}{G + e^{-\frac{AGB}{1200}}} \quad \textit{Eq.3}$$

From which the gross carbon gain follows an exponential increasing pattern with the biomass accumulation.

$$L = L_b + L_d \quad \textit{Eq.4}$$

The process of carbon loss consists of two parts: loss during disturbance events (L_d) and background mortality (L_b, Eq. 4). Here we specify that background mortality is a constant proportion of above-ground biomass, including the average effects of litterfall, root exudates, and herbivory (Thurner et al. 2016).

$$L_b = AGB \times K_b \quad \textit{Eq.5}$$

$$K_b = \frac{1}{\tau} \quad \textit{Eq.6}$$

Where K_b refers to the background mortality rate, which is the reciprocal of the turnover rate τ. On the other hand, disturbance-induced carbon loss occurs as a result of disturbance events, and the proportion would be determined by the intensity of the specific disturbance event which covers the patch.

$$L_d = AGB \times I \quad \textit{Eq.7}$$

Parameter I represents the disturbance's intensity, depending on the size of the event and the intensity slope, which will be addressed below in Section 2.2.

2.2 Model different disturbance regimes

We applied three parameters to describe the three dimensions of disturbance regimes: μ, α, and β, representing the probability of total occurrence, clustering pattern of events, and intensity slope respectively in the domain. For the purpose of distributing a sufficiently large and spatially random number of disturbance events, our experiment deploys a million-patch array to simulate the corresponding landscape-scale domains. Each patch represents one single canopy tree square with a 10 m length, so the total domain size is 10,000 ha or 100 km^2.

2.2.1 Parameterization of disturbance regimes

Specifically, the disturbance parameter μ refers to the total disturbed area in the domain:

$$D_a = D \times \mu \quad \textit{Eq.8}$$

Here D refers to the total domain size and μ is the percentage of the domain affected by disturbances.

As for parameter α, we followed Fisher's method by applying the scaling exponent α to determine the clustering degree of events (Fisher et al. 2008). The lower α represents the disturbance events would be more clustered, exhibiting the characteristics of large disturbing events but with rare occurrences in the domain; in contrast, the higher α will distribute the total disturbed areas more scattered and simultaneously increase the occurrences of small-sized events in the domain. To simulate multiple scaled events, the event sizes are prescribed as a numerically discrete series from 2^0 cells to the domain size, capped at 2^{19} cells. Stepwise values follow the mechanism of powers of 2, leading to twenty classes of event sizes:

$$n_z = Az^{-\alpha} \quad \textit{Eq.9}$$

Where the number of disturbance events (n_z) at a specific event size z is following a power law mechanism, α is the scaling exponent for the disturbance event clustering degree with a dimensionless unit. And the value A is a proportionality factor, manipulated by the size of the total disturbed area and the setting of events size series, and z_i is the corresponding disturbance event size, maximum i is 20 in this case:

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

Due to the discrete nature of the event sizes (z_i) and the pseudo-random amount of the corresponding events, a limited uncertain gap remains between the total disturbed area after the generation process and the prescribed value. In an attempt to limit this gap, we performed an error threshold to regulate this randomness: the difference between generated total area and prescribed value as a percentage of total domain area should be lower than 0.001% (10 pixels to a 1000-width domain). When the gap has exceeded the threshold, the new event amount sequence will be recalculated until the condition is met. In very rare cases, it is difficult to compute an amount sequence satisfying the threshold of 0.001%, so in which circumstances, the acceptable gap is relaxed to 0.002% (20 pixels to a 1000-width domain).

We assume the intensity of disturbance or the fractional mortality due to disturbance is proportional to the logarithm of event size (Chambers et al. 2013):

$$I = \beta \cdot log_{10}(Z_i) + b \quad \textit{Eq.11}$$

The parameter β controls the slopes of the logarithmic function for describing the relationship between the disturbance's intensity and its size. We descend from Chambers' description of the quantitative relationship between the average mortality rate and disturbance size (Chambers et al. 2013), inheriting a constant intercept parameter b = 0.22684 but varying interval of slope parameter β, from 0.03 to 0.5. For the same size of disturbance events, a larger β indicates a greater intensity, causing more carbon loss during the dynamic carbon cycling simulation. In

practice, it is possible for the value of intensity to exceed 1, which usually happens for the big beta and large events. In those cases, all intensity exceeding 1 should be limited back to 1 as per the reality of the situation.

2.2.2 Disturbance Parameter Ranges

The inclusion of the parameters μ, α, and β allows a flexible description of disturbance regimes in different scenarios of landscape, containing all possible disturbance properties in three dimensions and without the need for specifying particular types of disturbance. To be realistic, we referred to previous literature and designed different ranges and intervals for the primary productivity parameter G and three disturbance parameters independently.

In particular, the parameter G is specified as a range from 0.03 to 0.1, deriving a range of average steady state biomass from 10 to 40 $kgC \cdot m^{-2}$ without disturbances; the parameter μ is set in a range of 0.01 to 0.05 with an interval of 0.005, corresponding to the average documented value of 0.02 for forests (Moorcroft, Hurtt, and Pacala 2001; Malhi et al. 2004); α's setting comes from observed gap-size distributions from previous studies in the tropical and sub-humid forest ecosystem, which indicated a suggestive range of α within 1.1-1.6 (Lawton and Putz 1988; Jans et al. 1993; Nelson et al. 1994; Yavitt et al. 1995; Fisher et al. 2008), accordingly a 0.05 interval step was deployed in the range of 1.0 to 1.8 for the experiment; With regard to the parameter intensity slope - β, we refer to the statistics result from (Chambers et al. 2013)'s study and set the range from 0.03 to 0.5 with varied intervals (Chambers et al. 2013), which is 0.01 for the range [0.03, 0.09], and 0.05 for [0.1, 0.25] and 0.1 for [0.3, 0.5].

Table 1. Parameter Setting

Parameter	Range	Interval	Count
μ	[0.01:0.05]	0.005	9
α	[1.0:1.8]	0.05	17
β	[0.03:0.5]	0.01/0.05/0.1	14
G	[0.03:0.1]	0.01/0.02/0.03	5

2.3 Generation of Disturbance Events

To comply with the prescribed setting of disturbance regime parameters, a disturbance events generator was proposed to place all the events in a domain array, which is also supposed to satisfy the requirement of stochasticity in spatial distribution. For the convenience of calculations, all disturbance events are restricted to the shape of a rectangle (Fisher et al. 2008), the event lengths are specified as 10m, 300m, 500m, 750m, 1000m, and the corresponding width varies with the size of the disturbance events. The placement of disturbance events follows the principle of ascending from the largest to the smallest. The generator randomly assigns the center coordinates of a class of size disturbance events and detects whether they overlap with surrounding events that were already placed. Initiate

reassignment if overlapping is detected or continue with the assignment of the next class size disturbance events.

The disturbance events generator also includes a boundary overlapping detection function. If there are any events that intersect with the domain's boundary, leading to a partial disturbance that does not occur within the array, the generator will activate the correction response by reassigning the center coordinates for the overlapped disturbance events.

2.4 Characteristics of the Biomass Pattern

Following the previous parameter range and interval setting, 2142 combinations of disturbance regimes were established in total. For each regime, we constructed a cube of depth 100 as a reference database for the disturbance event. Every snapshot of the reference cube is a 1000x1000 stochastic disturbance reference map, providing the spatial reference for the domain's carbon cycling dynamic simulation.

With the support of the disturbance reference cube, we applied the strategy of unordered sampling with replacement to generate time series disturbance references. Ultimately, 200 maps of disturbance events were randomly extracted from cubes as a sequence of reference for simulating the occurrence of disturbances over 200 years in the domain. For each disturbance regime, we incorporated the disturbance sequence from the corresponding cube, together with the prescribed varied primary productivity capacity (parameter G), to run the dynamic carbon model to an equilibrium state of biomass. Motivated by the consideration of randomness in the occurrence of temporal disturbances, we shuffle the sequence of 200-year disturbance reference maps up to 10 times for each run of the model. The results of all runs will be labeled with the shuffle index and used for training or validation.

Despite some subtle sawtooth fluctuations that can instantaneously deviate from the expectation due to the impacts of stochastic disturbances, the average biomass for the whole domain saturates and reaches a dynamic steady state over the 200-year simulation run. We averaged the biomass maps for the last decade to obtain steady-state equilibrium biomass distributions, from which three kinds of statistical indicators were used to characterize steady-state biomass distribution properties under variable disturbance regimes and primary productivity.

The first type of statistics is based on the histogram distribution, such as mean, median, variance, skewness, kurtosis, percentiles, as well as standard deviation and coefficient of variation. Some of these metrics, such as skewness and mean, have been suggested to be associated with the probability and intensity of disturbances in previous literature (Williams, Hill, and Ryan 2013). We have also introduced a type of informative feature - the Shannon entropy index (Shannon 1948), to measure the "amount of information" inherent in the biomass map. Shannon's entropy, or the Shannon-Wiener index, is also a commonly used indicator that describes the diversity level in ecosystems (Spellerberg and Fedor 2003). In addition, the

statistical properties that come from the texture of biomass distribution have also been included. Texture provides information about the spatial arrangement of intensities in an image, a biomass map in our case, characterized by the spatial distribution of intensity levels in a neighborhood. One of the most common texture feature extraction methods based on image statistics is Gray-Level Co-Occurrence Matrices (GLCMs), which study the spatial correlation properties by using gray-scaled images (Haralick, Shanmugam, and Dinstein 1973). Here in our study, we applied four statistics from the GLCMs method: Contrast, Correlation, Energy, and Homogeneity (Table 2), to extract the texture features of the biomass maps. It is notable that all biomass maps are re-scaled to a range of 0 to 255 before calculation, and non-equilibrium outliers have been detected and removed.

Table 2. Statistics of the steady-state biomass map

Feature types	Statistic	Variable Names	Formula	
Histogram features	mean	mean_bio	$\frac{\sum_{i=1}^{N} A_i}{N}$	N: Total patch amount A_i: Biomass value for patch i
	median	median_bio	$Med(A)$	A: Biomass map
	range	range_bio	$P_{90} - P_{10}$	P90: Percentile 90% P10: Percentile 10%
	variance	var_bio	$\frac{1}{N-1}\sum_{i=1}^{N}\lvert A_i - \mu\rvert^2$	μ: Mean biomass
	standard deviation	std_bio	$\sqrt{\frac{1}{N-1}\sum_{i=1}^{N}\lvert A_i - \mu\rvert^2}$	
	coefficient of variation	cv_bio	$100 \times \frac{\sigma}{\mu}$	σ: Standard deviation
	skewness	skew_bio	$\frac{\sum_{i=1}^{N}(A_i - \mu)^3}{(N-1)\sigma^3}$	
	kurtosis	kurt_bio	$\frac{\sum_{i=1}^{N}(A_i - \mu)^4}{(N-1)\sigma^4}$	
	percentile 25%	prc25_bio	P_{25}	P25: Percentile 25%
	percentile 75%	prc75_bio	P_{75}	P75: Percentile 75%
	Trimean	Trimean	$(P_{25} + 2 \times MED + P_{75})/4$	MED: Median value
Informative feature	Shannon entropy	Shannon	$-\sum p \cdot log_2(p)$	p: Normalized histogram counts
Texture features	contrast	contrast	$\sum_{i,j}\lvert i-j\rvert^2 glc(i,j)$	i: Reference pixel value j: Neighbor pixel value glc(i,j): An entry in GLCM μ_i, μ_j: Means of GLCM w.r.t. i and j σ_i, σ_j: Standard deviations of GLCM w.r.t. i and j
	correlation	correlation	$\sum_{i,j}\frac{(i-\mu_i)(j-\mu_j)glc(i,j)}{\sigma_i\sigma_j}$	

	energy	energy	$\sum_{i,j} glc(i,j)^2$
	homogeneity	homogeneity	$\sum_{i,j} \frac{glc(i,j)}{1+\lvert i-j \rvert}$

As mentioned above, we have assumed 2142 different combinations of three disturbance regime parameters, and each individual regime corresponds to 5 different primary productivity scenarios. Therefore, more than ten thousand parameter combinations have been involved in the simulation of biomass dynamics with the carbon cycling model, and each parameter combination is expected to run ten times via different shuffling sequences of disturbance reference maps. The simulation will produce 17 statistics in total for each run, 16 comes from the steady-state biomass distribution and 1 is the mean GPP at the end of the simulation.

2.5 Prediction and validation

We performed a multi-output random forest regression method by Scikit-learn (Pedregosa et al. 2011), to investigate the links between the statistics from simulated biomass maps and the prescribed disturbance regimes and primary productivity parameters. In order to ensure the validity of the trained model and avoid overfitting, we designed three types of cross-validation strategies.

With the prescribed disturbance regime and GPP parameters, all the steady-state biomass distribution statistical metrics were documented together with the disturbance sequence shuffle index. For the first type of cross-validation, referred to as the Completely Random 10-fold method, we disrupted all entries in random order and equally divided them into ten parts for 10-fold cross-validation. Nine-tenths of the data is used to fit the model and the rest is for validation, and the ultimate prediction accuracy is the mean of ten cross-validation results. Another strategy refers to the Leave-One-Sequence-Out method (LOSO). Instead of randomly dividing all the data into ten sets, the fit and validate process was conducted according to the shuffle index. For instance, entries with shuffle index 1 were used for validation and the rest for training and circulated the validation shuffle index until all the data are validated and trained. The last strategy is against the robustness of disturbance regime parameters. For each of α, β, and μ, we keep each value in turn for validation and train all the remaining data to test the model's predictive capacity for the untrained parameters.

We use Nash-Sutcliffe efficiency (NSE), to evaluate the performance of the predicting model (Nash and Sutcliffe 1970). The bigger NSE represents the better accuracy between prediction and observation, which are predicted and prescribed disturbance regime parameters in this study.

$$NSE = 1 - \frac{\sum_{i=1}^{n}(Y_i^{obs} - Y_i^{pre})^2}{\sum_{i=1}^{n}(Y_i^{obs} - \overline{Y^{obs}})^2} \qquad \textit{Eq.12}$$

Where the Y_i^{obs} is the i-th prescribed disturbance regime parameter, Y_i^{pre} is the i-th predicted value for the constituent being evaluated, $\overline{Y^{obs}}$ is the mean of the prescribed parameter, and n is the total number of observations (D. N. Moriasi et al. 2007).

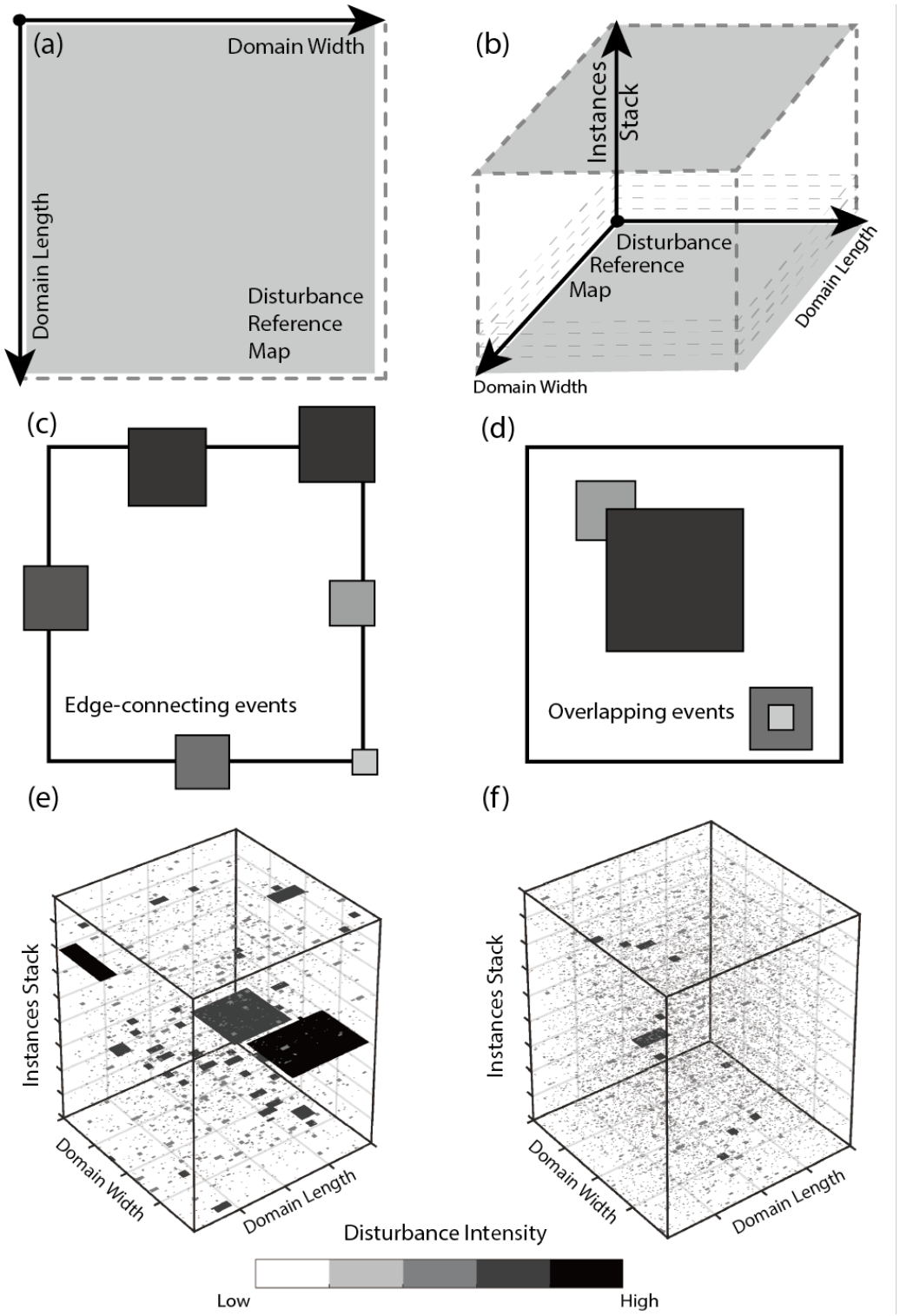
3 Results

3.1 Disturbance reference cubes

3.1.1 Spatial binary cubes

The disturbance generator produced 153 (9 *μ*'s and 17 *α*'s) disturbance binary reference cubes in total, each cube representing a unique combination of *μ* and *α* and consisting of 200 snapshots simulating different scenarios of different disturbance event spatial distribution. Notably, these snapshots are all binary and without any attribute related to the information of the intensity, which means that they provide only the spatial reference information under specific mu and alpha.

The disturbance events, which refer to the unique patches with true value, regardless of their corresponding sizes or shapes, are meant to be randomly distributed across the whole domain without overlapping or overstepping boundaries. As two cubes with completely different disturbance regimes (Figure 1e-f), all the disturbance events are uniformly distributed throughout the whole domain, satisfying the requirement of no overlap and relative independence (Figure 1c-d), despite the fact that there are cases where large events are joined by some small events. Other than that, all marginal areas are well circumvented, thus ensuring that no pixels are omitted, otherwise causing the simulated *μ* to be less than the specified value.



365 *Figure 1. Concept diagram of disturbance reference cubes, (a) refers to the 2D disturbance reference map, (b) refers to the 3D disturbance reference cube, and (c) disturbance events should avoid interfacing with edges or (d) overlapping with each other. (e) and (f) show two disturbance reference cube examples with different disturbance regimes.*

370

3.1.2 Intensity reference cube

Intensity of disturbance, or the fractional carbon loss due to disturbance, is designed of being varied with event sizes by parameter β. As a consequence of the binary spatial reference cube, the varied β, which represents the gradient relationship between event size and intensity, were assigned to the same cube and resulted in different intensities.

3.2 Parameterization and dynamic carbon model

3.2.1 parameterization of disturbance regimes

Parameter α ensures that the disturbance event in a specific size and its amounts across the domain consistently adhere to a binding correspondence. Take an example of α-1.2 (Fig S1), the number of events decreases exponentially with the event size. Although the theoretical largest event has an area of up to half the size of the total domain (2^{19} pixels), the maximum event size in this experiment is limited to 2^{14} due to the range of alpha parameters (Table1, minimum alpha is 1.0).

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 concentrated large ones. With discrete but progressive alpha values, the disturbance reference cubes provide a gradient relationship between the amounts and sizes of events across the domain.

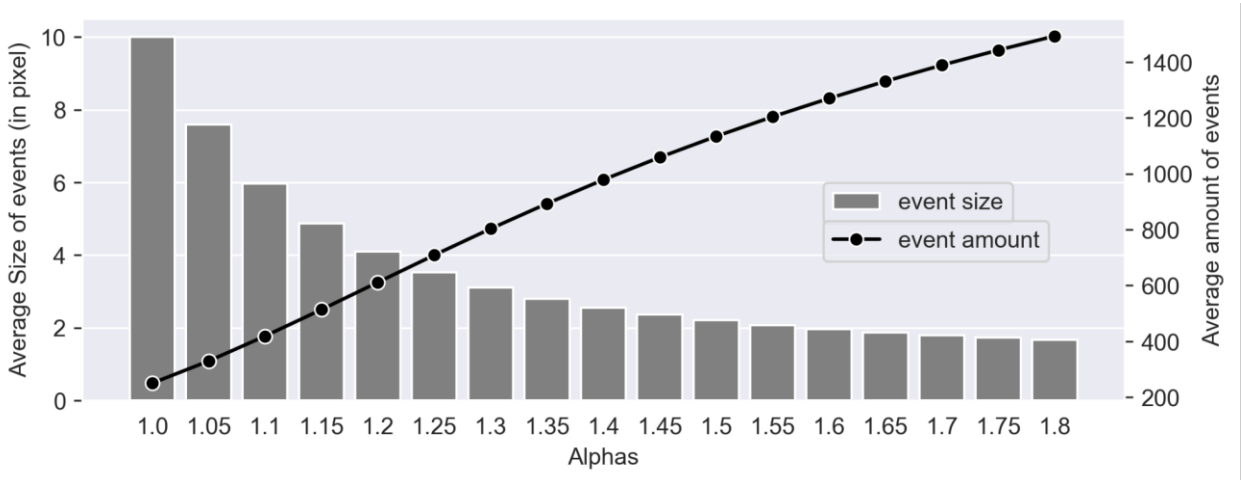


Figure 2. The relationship between gradient α and the average sizes of events (bar plot) and the average amount of events (line plot) in domain.

Parameter beta, designed in a hierarchy gradient, is in an attempt to lead to different intensity levels with the same event size. According to the prescribed relationship between β and event size (Eq.9), the intensity of disturbance events needs to follow a logarithmic pattern

with its increasing sizes. Figure 3 shows this pattern of event size and intensity for all the betas we assigned in the experiment, and two different scales for small-size events (linear x-axis, range 1-32) and larger ones (logarithmic x-axis, range 32+) were employed to bring out the regularity of this relationship.

The result shows all curves maintain a clear hierarchy between each other, indicating the gradient relationship between intensity and event sizes. For small-size events, they share the same instant logarithmic increment for intensity, and then the intensity goes much slower with the increasing event sizes. *β* greater than 0.2 are likely to exceed the intensity greater than 1, in which case the intensity would be limited to 1, corresponding to reality. Notably, *β*-0.5 reaches the saturation of intensity limitation at the very beginning of event size, which performs a straight line on the top of the plot.

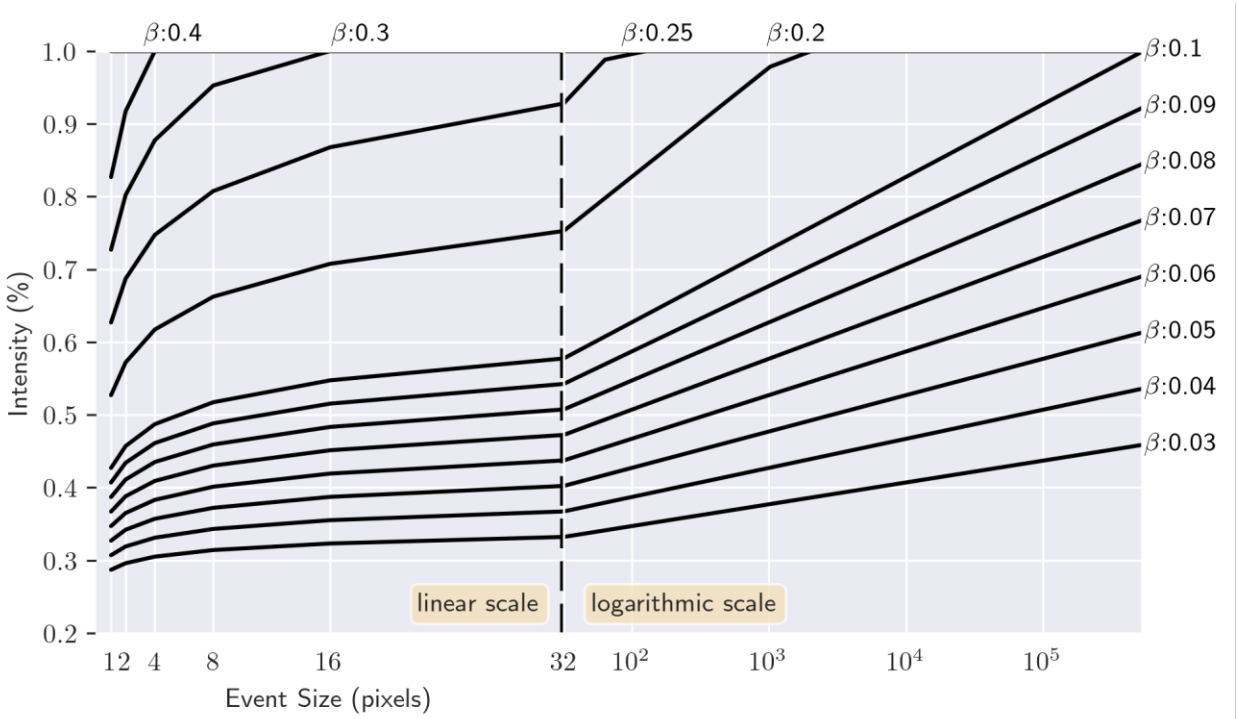


Figure 3. Relationship between Intensity and event size, showing linear scale for events under 32 pixels and logarithmic scale for larger events. The intensity value of β-0.5 is saturated at 1.0 starting from event size 1.

3.2.2 Temporal carbon dynamics

Using the parsimonious carbon model, the dynamics of GPP together with carbon loss from physiological mortality and disturbances, can be simulated in an annual step and lead to a logarithmic growth for biomass. The parameters G, designed to represent gradients of photosynthesis, produce hierarchical GPP capabilities at steady states without taking disturbances into account. Meanwhile, AGB reaches different steady states with five different Gs, indicating different scenarios with significant differences in average biomass across

domains (~ 10 $kgC \cdot m^{-2}$ to 40 $kgC \cdot m^{-2}$). Figure 4 also shows the impacts of disturbance on temporal GPP and AGB dynamic. In the case of a disturbance regime with *μ* 0.04, *α* 1.5, and *β* 0.5, although the overall growth pattern has not changed much, the average level of GPP and AGB in steady states are significantly reduced due to the involvement of disturbance, GPP reduced from ~0.5 to 1.5 $kg \cdot m^{-2}$ for different Gs, and AGB dropped to half of its original level on average. Furthermore, the original smooth growth curves have been replaced with distinct sawtooth fluctuations that become more violent as G increases.

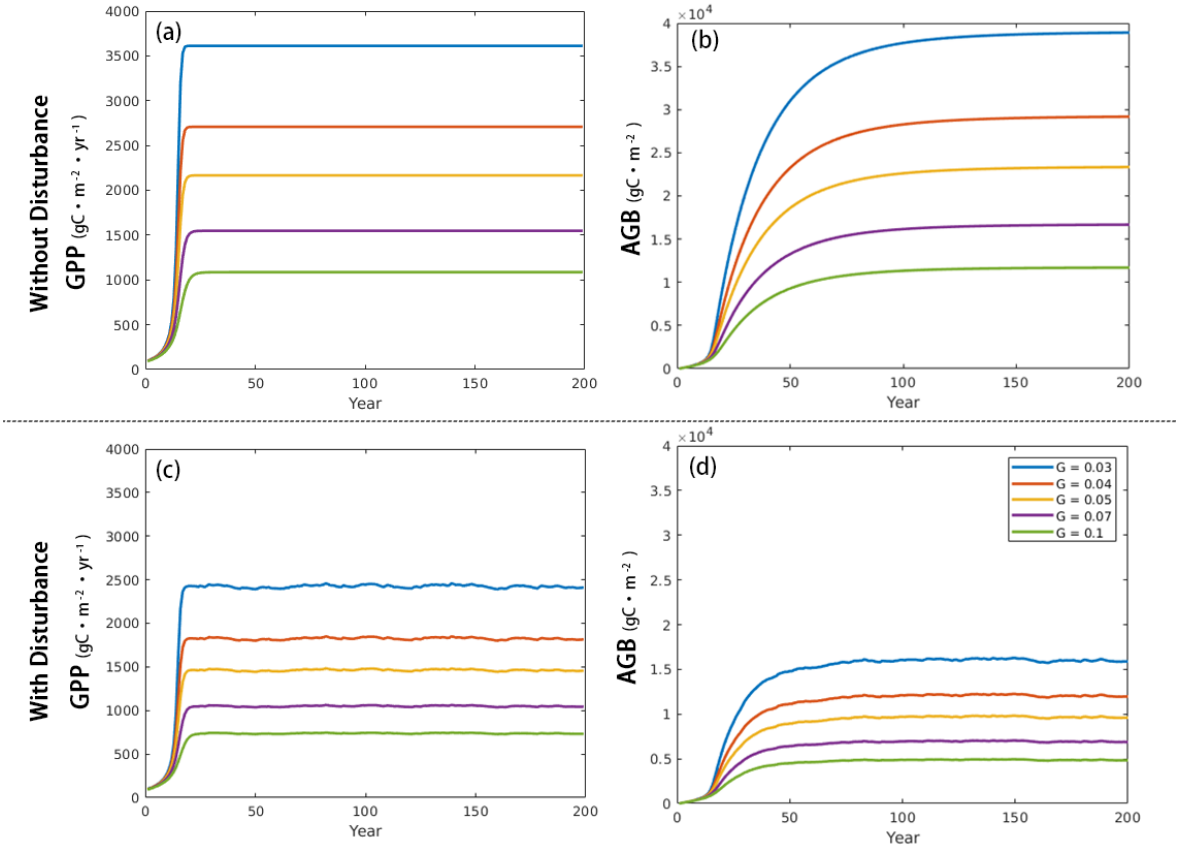


Figure 4. AGB and GPP evolution with varied Parameter G, (a) and (b) refer to the scenario without any disturbance events, (c) and (d) provide the impact of disturbance regime with μ-0.03, α-1.0, and β-0.2.

3.2.3 Steady-state Biomass distribution

Under the same capacity of photosynthesis (fixed G), the steady-state biomass (year 200) shows diverse spatial and numerical differences with different disturbance regimes. By increasing parameter *μ* and fixing the other two parameters (first row in Figure 5), there are more areas with lower biomass, which exhibit more signs of impact from disturbances and indicate the increasing possibility of total disturbance occurrences. In addition, as the level of clustering degree changes, at which the parameter *α* increases, the spatial distribution of the steady-state biomass map becomes more homogeneous, shifting from being able to distinguish

the effects of disturbance events to a uniformly distributed noise mosaic (second row in Figure 5). Furthermore, the influence of elevated β is apparent as well. Events with higher intensity removed more biomass and left the succession of plants regrowing at a lower level, which results in a more clearly visible disturbance footprint. The combination of three disturbance regime parameters resulted in varied patterns of spatial biomass distribution, implicating a plausible possibility to retrieve disturbance attributes back according to biomass observation and dynamic biomass simulation model.

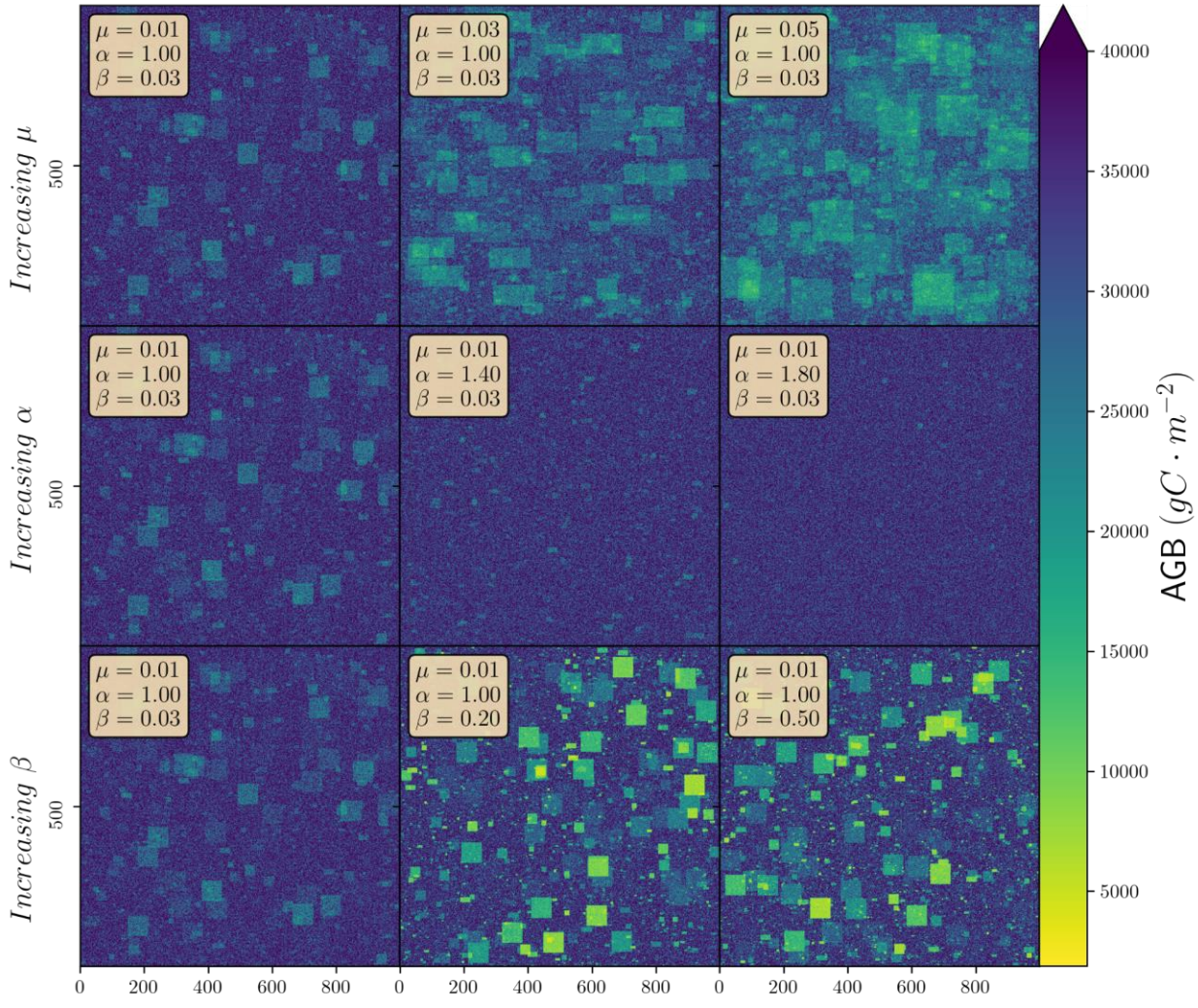


Figure 5. Steady-state biomass map in different disturbance regimes. The first row shows the comparison of increasing μ, the second row refers to increasing α, and the third row shows increasing β.

3.3 Retrieving Disturbance Regimes from Biomass Patterns

3.3.1 Cross-Validation

Table 3 shows that all the strategies of cross-validation have a good performance for retrieving the three disturbance regime parameters. Relying on the 16 statistics (table 2)

calculated from the average biomass of the last ten-year simulation and the steady-state GPP constraints, all the NSEs of μ, α, and β exceeded 0.95 (Table 3) in the CR and LOSO validation. In addition, the LOPO validation is also relatively accurate, in the case that the target prediction parameters are not involved in the training, they can still be predicted with the NSE around 0.88.

Table 3. Cross-Validation results via different strategies

Average NSE	Completely Random 10-Fold (CR)	Leave One Sequence Out (LOSO)	Leave One Parameter Out (LOPO)
μ	0.973	0.971	0.881
α	0.966	0.964	0.878
β	0.979	0.978	0.882

The scatter plots of different CV strategies confirmed the high accuracy and apparent gradient for the predictions. Figure 6a is an example of the completely random prediction results, for μ and α, equally spaced parameter settings correspond to a relatively uniform distribution outcome for each of the values, where the high-density pattern lies near the prescribed value. Only very few of the predicted values are discrete anomalies, whose amount does not exceed 10. Although the distribution is more scattered for the higher β values, due to the setting of varied intervals, the trend is also in line with the 1:1 line, indicating the great accuracy and high correlation with the prescribed values. In contrast to the CR validation, the LOSO validation (Figure 6b) maintains a similar prediction accuracy but the LOPO results are much more scattered. This may indicate that the accuracy of prediction is more sensitive to parameter interval setting rather than the shuffle order, particularly missing boundary values of the parameter during the training could lead to an obvious bias, such as the prediction of μ and β, presenting an integral shift to the inner domain for the maximum and minimum parameters.

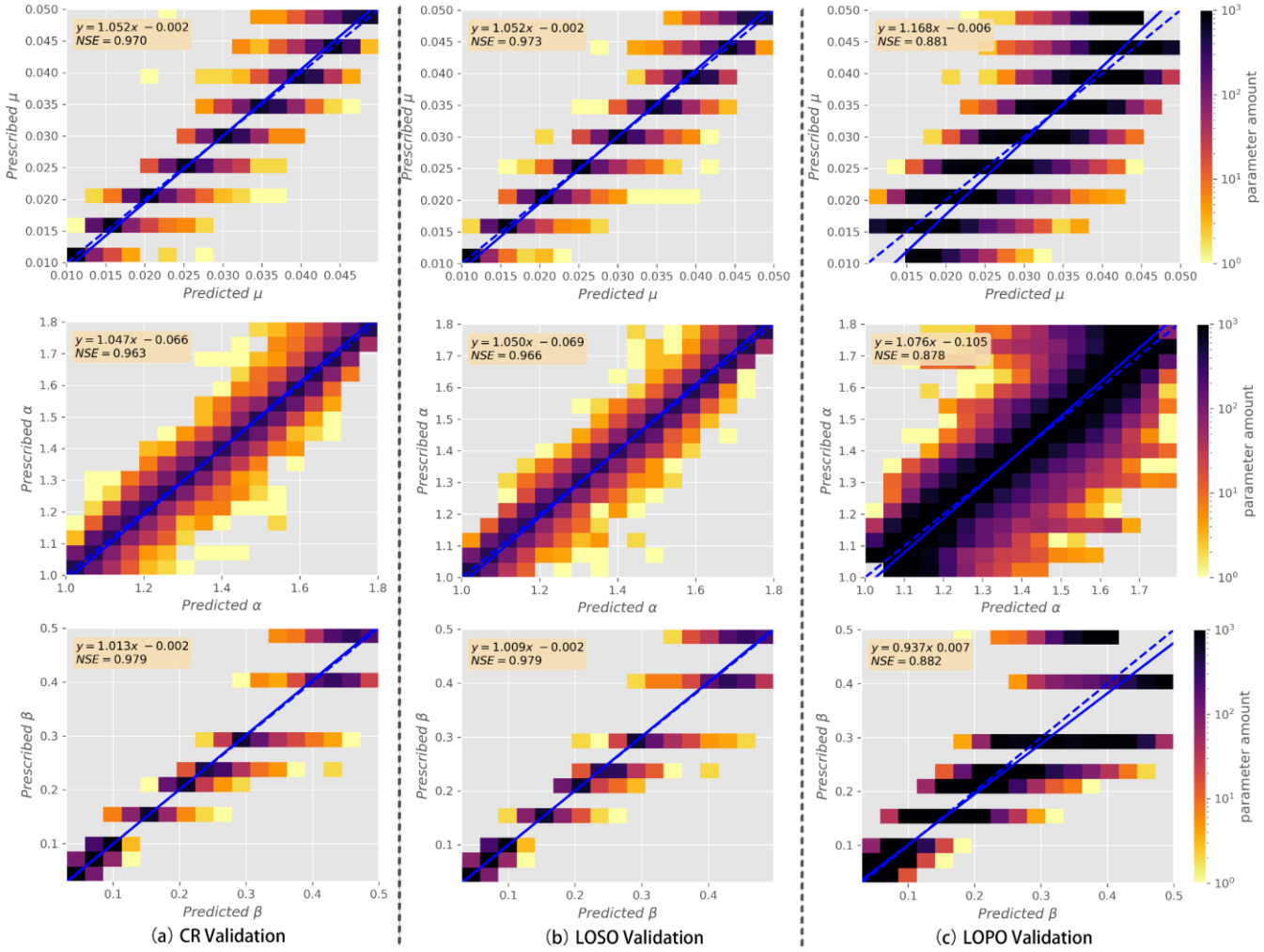


Figure 6. Accuracy of the completely random validation for three disturbance regime parameters. The horizontal axis denotes the predicted values and the vertical axis denotes the prescribed values. (a) illustrates the three-parameter prediction results in the Completely Random cross-validation strategy, (b) refers to the Leave One Sequence Out strategy, and (c) refers to the Leave One Parameter Out strategy

3.3.2 Feature Importance

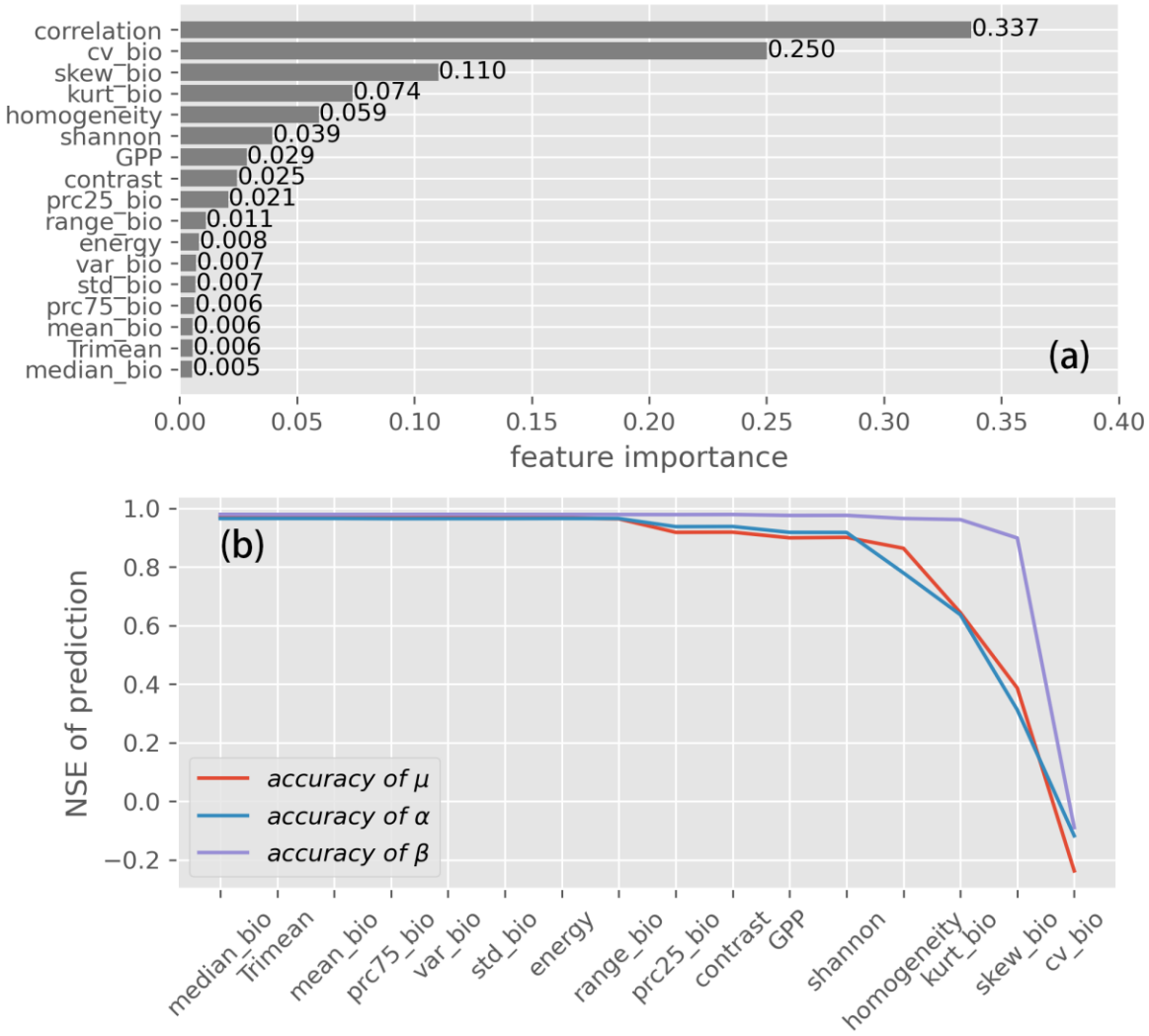


Figure 7. (a) shows the feature importance of multi-output disturbance regime prediction, the value for each feature represents the magnitude of the accuracy decrease when removing its variability from the prediction of all three disturbance regime parameters; and (b) shows the prediction accuracy change in ascending importance for each disturbance regime parameter, x-axis refers to the feature(s) excluded from the prediction process (cumulatively from left to right). Accuracy is measured as the NSE between the full model estimates and the model outputs from the different factorial experiments for the entire three disturbance regime metrics (μ, α, β).

In the multi-output random forest method, the feature importance analysis reveals that all types of statistics played significant roles in predicting the disturbance remiges, the information of steady-state GPP constraints has around 3% contribution (Figure 7a), ranking at seventh place. More specifically, the texture feature correlation, together with the histogram feature coefficient of variation, dominates the main contribution to prediction, with which about 60% of the predicting contribution for disturbance regime parameters can be explained.

We also implement an exclusion experiment to quantify the prediction accuracy of μ, α, and β by removing one feature each time, following the order of contribution from lowest to second

highest. Figure 7b shows that both μ and α start with a steep drop in prediction accuracy from the Shannon index (6th in the importance order), while β doesn't change significantly in NSE until only the last two features (correlation and coefficient of variation).

In order to investigate the relationship between the individual disturbance regime parameter and statistics, we retrained the corresponding random forest model for only predicting one parameter by using biomass statistics and GPP constraints. Figure 8a illustrates that μ is mainly controlled by histogram features, and GPP ranked fourth with exceeding 10% contribution, playing the most significant role compared to α (< 4%) and β (<2%). Texture features are the main predictors for parameter α, which can explain about 50% contribution of the prediction. The largest contributor-correlation accounts for about 40%. In the case of β, 80% of its contribution is due to the coefficient of variation, which occupies an absolute position.

A similar exclusive analysis was applied to the individual parameter, and the result shows that to achieve over 80% accuracy, μ, α, and β require the participation of at least 2,3, and 4 features respectively (Figure e-g).

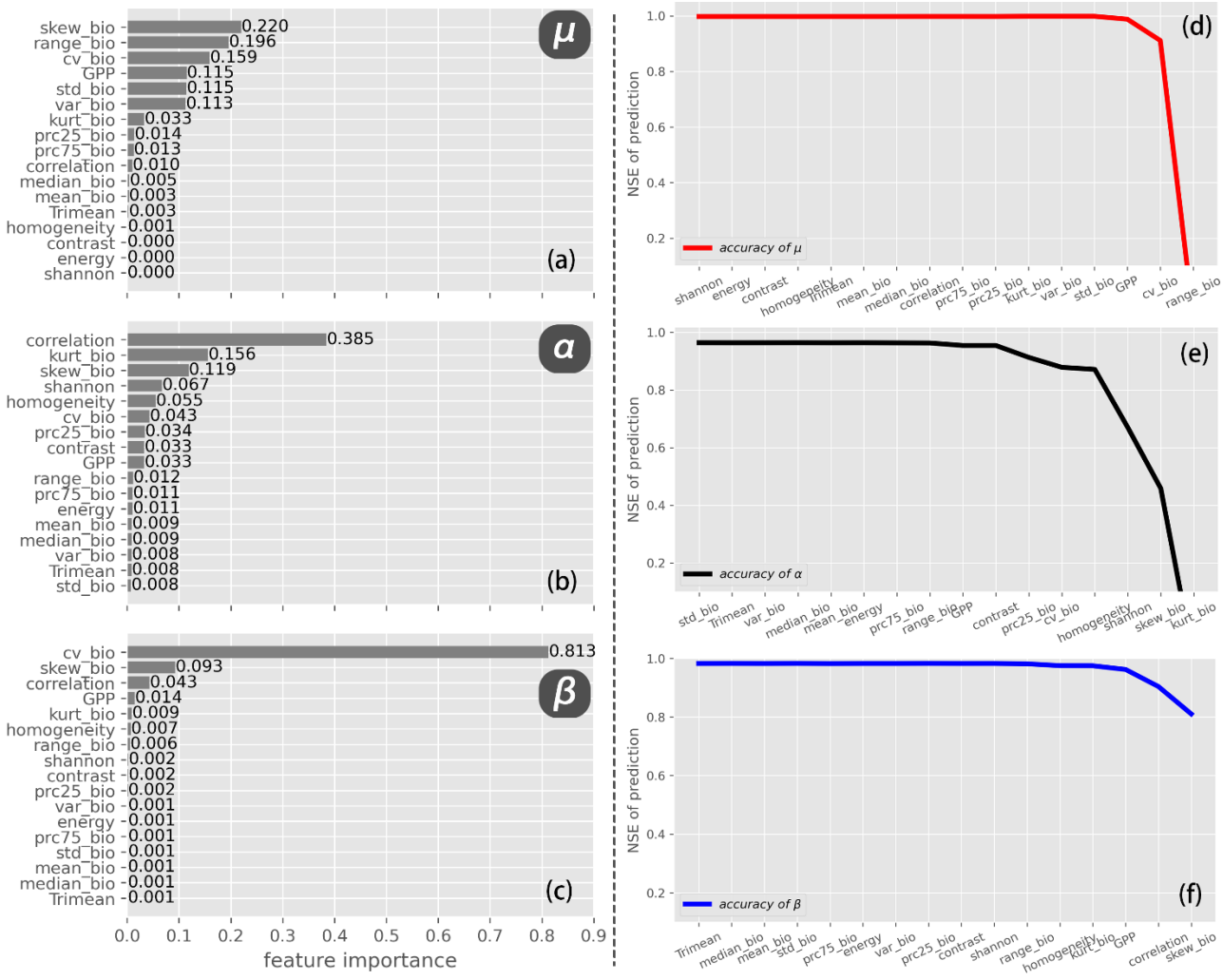


Figure 8. Breaks down the results of the feature importance for each individual disturbance regime parameter: μ, α, β. On the left (a-c) are shown the feature importance results (as in Fig. 7a) while on the right (d-f) the cumulative feature importance (as in Fig. 7b).

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). Some researchers have attempted to infer or evaluate partial representatives of the disturbance regimes in different manners from the model and observation, such as the number, size, and distribution of disturbance events in Fisher's work (Fisher et al. 2008), probability and intensity in Williams' work (Williams, Hill, and Ryan 2013), or return frequency, fractional trees and fractional mortality in Chambers' work (Chambers et al. 2013). Based on these previous studies as well as the definition, we designed three general parameters μ, α, and β with a relatively wide parameter range to simulate sufficient varied disturbance regime scenarios without considering the specified types. In terms of parameterization, the binary disturbance reference maps are controlled by μ, α which are following Fisher's work (Fisher et al. 2008), these two parameters were employed at equal intervals for conducting gradient patterns of spatial distribution for disturbance events. From the prediction results, the evenly distributed and distinguishable μ and α show that the set of intervals is reasonable and indicates the capability of a denser improvement (Figure 6). Instead of killing all of the biomass due to any disturbance events, we use parameter β to link the disturbance intensity or the fraction of biomass loss with the corresponding event size. Based on the statistic from Chambers' work, we experimentally employed a set of values around the statistic with varying intervals, and the prediction shows a more divergent prediction distribution but also exhibits good prediction accuracy as the intervals increase. As a result, we suggest that the modelers focus equally on the interval and range, and the limits and validity of settings should in our study be evaluated further.

Here we hypothesize and demonstrate that higher-order statistics on biomass patterns reveal properties of underlying disturbance regimes which have been assumed to be unsolvable in previous modeling exercises (Fisher et al. 2008; Ryan et al. 2012; Williams, Hill, and Ryan 2013). We present a conceptual modeling framework where a wide variety of disturbance regimes is inverted from simulated biomass patterns via machine learning. Our results are robust for different cross-validation strategies, suggesting the generalization of the approach (Table 3).

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, on particular disturbance metrics, the rankings of variable importance may change. Mu, the average domain fraction affected by disturbances, is strongly linked to biomass distribution statistics (skewness, range, CV, standard deviation) and also to the GPP (Figure 8a). Interestingly biomass itself does not emerge as of high importance, although it is implicitly present in CV and STD. Alpha, the clustering pattern of events' distribution, is more correlated with text features, such as Correlation, Homogeneity, and

Contrast, which account for around 50% of contributions (Figure 8b). And Beta, controlling the relationship between disturbance intensity and event size, is mainly dominated by CV (Figure 8c) whose contribution is over 80%. This connection is most obvious among the three parameters, indicating that the intensity would directly affect the biomass histogram distribution and thus lead to varying CV statistics.

By design, within each domain, (1) disturbance events are independent between years, and (2) the event shape does not change according to event type or local land surface conditions, while (3) GPP parameters do not change within the domain.

(1) As the temporal occurrence, the current setting of disturbance events are all temporarily stochastic for considering the simplicity of simulation. We randomly extract a single snapshot annually from the disturbance cube, in which the instances within the stack are also independent. In reality disturbance, however, could modify forest soil's physical and chemical factors and microclimatic environments, creating ecological legacies that have further legacy effects on carbon dynamics (Liu et al. 2003). Thus different disturbance types would result in a varying probability for the occurrence of the subsequent disturbance event. Further discussion is needed to constrain this temporal randomness with a view to narrowing the gap between reality and simulation.

(2) As for the event shape, all the disturbance events are regulated as rectangular shapes for the reason of computing efficiency at the moment. This strategy allows all the events across the domain can be distributed without overlapping or edge-connecting with impressive efficiency. However, this simplified method could also raise a few problems. Firstly, the gaps between the rectangular events over time are likely to be missed even after a 200-year spatial-stochastic simulation, resulting in some extremely large values in non-steady states. These non-equilibrium outliers, defined as values more than three standard deviations from the mean value by each column of the domain (Table S2), were filled with the nearest steady-state values around. Although increasing the spin-up time may theoretically mitigate this non-equilibrium phenomenon, it's not cost-effective in terms of computation consumption. We also noticed that the statistical features without outliers preprocessing could lead to an apparent decrease in accuracy by a comparison experiment, this suggests that the correlation between biomass statistics and disturbance regimes relies on the steady-state assumption, and the noise of non-equilibrium outliers can significantly affect the prediction accuracy. Secondly, the shapes of realistic disturbance events are usually more complex and stochastic based on the monitor results of high-resolution remote sensing observation (Chambers et al. 2013; Senf and Seidl 2021a). The impact of varied disturbance event shapes on the statistical features of spatial steady-state biomass, and whether the simplification process will have any effect on the robustness of the prediction accuracy both need to be further discussed.

(3) For GPP parameter G, we applied 5 different values for representing a gradient photosynthesis capacity at the landscape level. Considering the heterogeneity of

photosynthesis capacity within the domain, we deployed a normal distribution to randomly distribute G values for each patch, but overall Gs did not vary much and remained fixed over time. It has been found that multiple environmental factors, such as temperature, vapor pressure deficit, atmospheric CO_2 concentrations, soil water availability, light intensity, and cloudiness, dominate varied impacts on the light use efficiency (LUE) model parameters across in-situ sites (Bao et al. 2022). Therefore more environmental variables and photosynthetic parameters should be taken into account to localize the GPP dynamic simulation.

Besides, we employed a relatively simple carbon dynamic model to simulate the dynamics of biomass, particularly the carbon gain and physiological loss mass are only associated with the plant biomass at the moment, exclusively allowing the analysis of the direct disturbance impact without considering other detailed physiological processes and allocation mechanisms. However, the realistic biomass dynamic is often more sophisticated, most of the compartment models at the moment, such as CASA (Potter et al. 1993), JSBACH (Reick et al. 2021), DALEC (Williams et al. 2005), etc., are capable of describing different biomass compartments such as leaves, branches, stems, roots, and CWD in various details. Besides the plant growth recovery can be affected by the climate and atmosphere in terms of long life span, the impacts of the changes like carbon dioxide fertilization and arising temperate on forests should be addressed (Norby et al. 2001; Pan et al. 2010). Therefore performing a more detailed, synthetic, and goal-directed model is expected, relying on which the direction of assessing the weights of these biomass-leading factors in predicting the disturbance regimes could be pointed and thus enable to evaluate of how these factors' interaction impact the ultimate prediction accuracy. Besides, through the incorporation of multiple observed constraints, more detailed carbon compartments and plant physiological processes like metabolic activity can be described, permitting the localization of regional biomass dynamics and corresponding disturbance regimes.

Identifying disturbances' varying occurrences in time and space and the complexity of the plant's response in different compartments remains a key scientific challenge. Through the use of remote sensing data and ground-based networks, significant advances have been achieved in understanding, representing, scaling, and characterizing disturbance, 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 (Parton et al. 1987; Running and Gower 1991; Raich et al. 1991; McGuire et al. 1992; Chen, Chen, and Cihlar 2000; Liu et al. 2003; Bond-Lamberty et al. 2005), 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 and Cienciala 2006; 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 (Shugart, Leemans, and Bonan 1992; Hurtt et al. 1998; Bugmann 2001; Norby et al. 2001). Overall, most of these processed-based models at the moment rely on the field or satellite observations to quantify or evaluate the impacts of disturbance on carbon stocks or fluxes, so the models are more diagnostic and require specific

information, such as the extent, type, and timing of disturbance events for supporting the impact analysis. To determine the impact of disturbance, the timely and continuous land change land cover (LULC) or other types of datasets describing the disturbance attribute always need to be taken into consideration in order to generate a more robust simulation of carbon dynamics (Liu et al. 2003). Based on the framework we built, the concept of the demography model simulating the regional biomass with replicated patch pixels and the mechanism of the carbon compartment model describing the biomass dynamic can be combined together to conduct the simulation. In spite of the dependence on a large amount of computation-consuming resources, our study is capable of retrieving the disturbance regimes parameter back only by the biomass observations under the assumption of the steady-state, empowered with prognosis for the development of many dynamic global vegetation models (DGVMs), such as LPJ (Haxeltine and Prentice 1996), HYBRID (A. D. Friend et al. 1997), IBIS (Foley et al. 1996), VECODE (Brovkin, Ganopolski, and Svirezhev 1997), and LM3V (Shevliakova et al. 2009) considering the impact of disturbance.

5 Conclusion

Our study reveals a strong link between landscape-level disturbance regimes and biomass distribution patterns and statistics. Relying on a conceptual model-based experiment and machine learning regression three disturbance regime parameters can be reliably retrieved via the spatial distribution of plant biomass and primary productivity. Namely, the average fraction of the domain affected by disturbances, the event size clustering exponent, and the perturbation intensity can be determined with an accuracy of 97.3%, 96.6%, and 97.9%, respectively. As earth observation efforts evolve to deploy upcoming global satellite missions such as GEDI (Stavros et al. 2017), NISAR (Rosen et al. 2015), and BIOMASS (Le Toan et al. 2011), further efforts to quantify photosynthesis at higher resolution are taking place (Cogliati et al. 2015; Jung et al. 2020), our approach lays the ground for the development of long term diagnostics on the terrestrial carbon cycle using contemporaneous observations.

Author contributions

NC and SW designed the experiments. In close collaboration with HY, SK, and NC, SW conducted the analysis and prepared the first draft of the manuscript. All authors contributed to the research discussions and improving the manuscript.

Competing interests

The authors declare that they have no conflict of interest.

Acknowledgement

Siyuan Wang acknowledges support from the International Max Planck Research School for Biogeochemical Cycles (IMPRS-gBGC). Siyuan Wang and Hui Yang acknowledge support from the German Federal Ministry of Economics and Technology (50EE1904).

Appendix

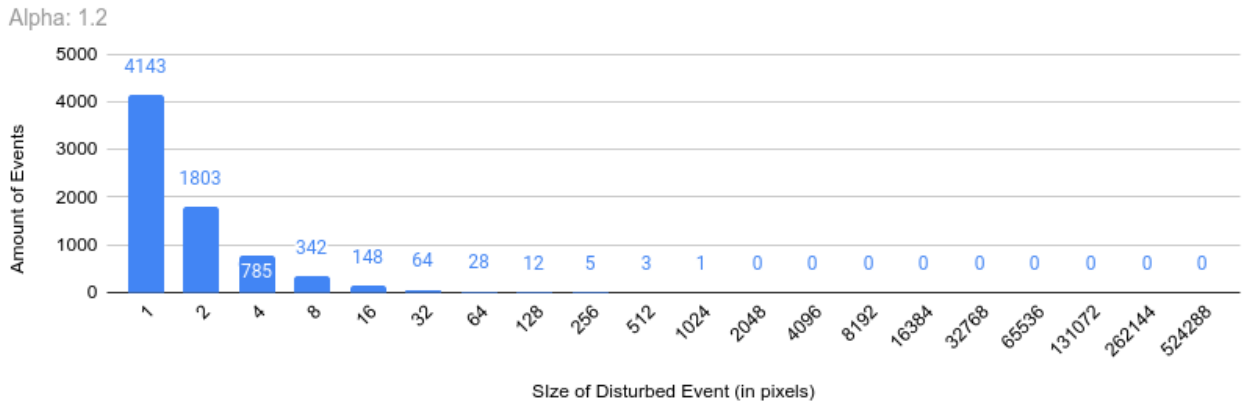
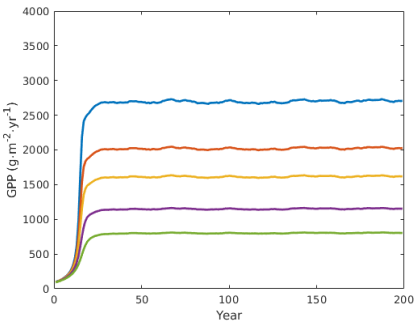
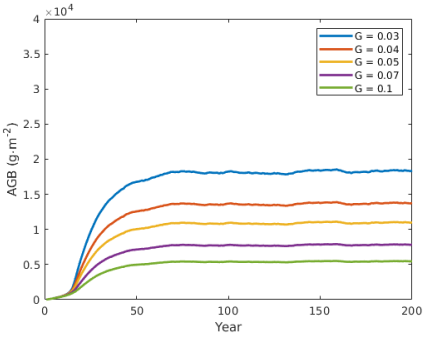
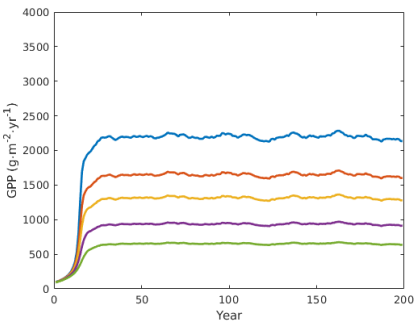
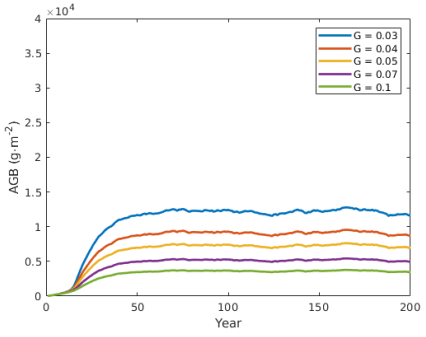


Figure S1. Number of events for each size in one domain where α is 1.2

	GPP	AGB
μ 0.03 α 1.0 β 0.2		
μ 0.05 α 1.0 β 0.2		

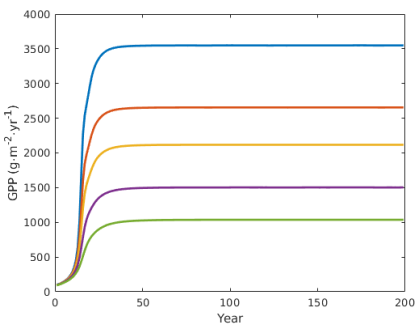
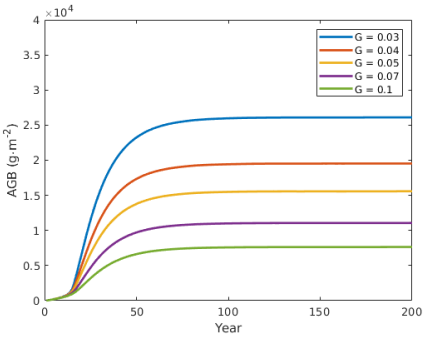
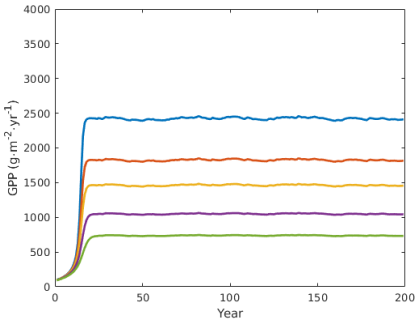
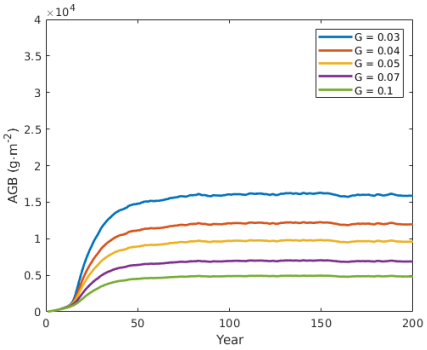
μ 0.03 **α 1.8** **β 0.2**		
μ 0.03 **α 1.0** **β 0.5**		

Figure S2. AGB ($gC \cdot m^{-2}$) and GPP ($gC \cdot m^{-2} \cdot yr^{-1}$) evolution against age (year) for different disturbance regimes

Table S2. the number of outliers in three methods for a domain with 1 million pixels

Pixels	Detect by each column	Detect by all values
Median	28602	28550
Mean	9887	9422
Quartiles	39628	34348

Reference

Allen, Craig, Alison Macalady, Dominique Bachelet, Nate McDowell, Michel Vennetier, Thomas Kitzberger, Andreas Rigling, 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 (January): 660–84. https://doi.org/10.1016/j.foreco.2009.09.001.

Amthor, Jeffrey 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.

Bao, Shanning, Thomas Wutzler, Sujan Koirala, Matthias Cuntz, Andreas Ibrom, Simon Besnard, Sophia Walther, et al. 2022. “Environment-Sensitivity Functions for Gross Primary Productivity in Light Use Efficiency Models.” *Agricultural and Forest Meteorology* 312 (January): 108708. https://doi.org/10.1016/j.agrformet.2021.108708.

Bond-Lamberty, Ben, Stith T. Gower, Douglas E. Ahl, and Peter E. Thornton. 2005. “Reimplementation of the Biome-BGC Model to Simulate Successional Change.” *Tree Physiology* 25 (4): 413–24. https://doi.org/10.1093/treephys/25.4.413.

Bossel, Hartmut, and Holger Krieger. 1994. “Simulation of Multi-Species Tropical Forest Dynamics Using a Vertically and Horizontally Structured Model.” *Forest Ecology and Management*, Contrasts between biologically-based process models and management-oriented growth and yield models, 69 (1): 123–44. https://doi.org/10.1016/0378-1127(94)90224-0.

Brokaw, Nicholas V. L., and Samuel M. Scheiner. 1989. “Species Composition in Gaps and Structure of a Tropical Forest.” *Ecology* 70 (3): 538–41. https://doi.org/10.2307/1940196.

Brovkin, Victor, Andrei Ganopolski, and Yuri Svirezhev. 1997. “A Continuous Climate-Vegetation Classification for Use in Climate-Biosphere Studies.” *Ecological Modelling* 101 (2): 251–61. https://doi.org/10.1016/S0304-3800(97)00049-5.

Brugnach, Marcela. 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, Harald. 2001. “A Review of Forest Gap Models.” *Climatic Change* 51 (3): 259–305. https://doi.org/10.1023/A:1012525626267.

Carvalhais, Nuno, Matthias Forkel, Myroslava Khomik, Jessica Bellarby, Martin Jung, Mirco Migliavacca, Mingquan Mu, et al. 2014. “Global Covariation of Carbon Turnover Times with Climate in Terrestrial Ecosystems.” *Nature* 514 (7521): 213–17. https://doi.org/10.1038/nature13731.

Chambers, Jeffrey Q., Niro Higuchi, Liliane M. Teixeira, Joaquim dos Santos, Susan G. Laurance, and Susan E. Trumbore. 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, Jeffrey Q., Robinson I. Negron-Juarez, Daniel Magnabosco Marra, Alan Di Vittorio, Joerg Tews, Dar Roberts, Gabriel H. P. M. Ribeiro, Susan E. Trumbore, and Niro Higuchi. 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): 3949–54. https://doi.org/10.1073/pnas.1202894110.

Chen, Wenjun, Jing Chen, and Josef Cihlar. 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.

Cogliati, S., W. Verhoef, S. Kraft, N. Sabater, L. Alonso, J. Vicent, J. Moreno, M. Drusch, and R. Colombo. 2015. "Retrieval of Sun-Induced Fluorescence Using Advanced Spectral Fitting Methods." *Remote Sensing of Environment* 169 (November): 344–57. https://doi.org/10.1016/j.rse.2015.08.022.

D. N. Moriasi, J. G. Arnold, M. W. Van Liew, R. L. Bingner, R. D. Harmel, and T. L. Veith. 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.

Delbart, N., P. Ciais, J. Chave, N. Viovy, Y. Malhi, and T. Le Toan. 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): 3027–39. https://doi.org/10.5194/bg-7-3027-2010.

Fisher, Jeremy I., George C. Hurtt, R. Quinn Thomas, and Jeffrey Q. Chambers. 2008. "Clustered Disturbances Lead to Bias in Large-Scale Estimates Based on Forest Sample Plots." *Ecology Letters* 11 (6): 554–63. https://doi.org/10.1111/j.1461-0248.2008.01169.x.

Foley, Jonathan A., I. Colin Prentice, Navin Ramankutty, Samuel Levis, David Pollard, Steven Sitch, and Alex Haxeltine. 1996. "An Integrated Biosphere Model of Land Surface Processes, Terrestrial Carbon Balance, and Vegetation Dynamics." *Global Biogeochemical Cycles* 10 (4): 603–28. https://doi.org/10.1029/96GB02692.

Franklin, Jerry F., H. H. Shugart, and Mark E. Harmon. 1987. "Tree Death as an Ecological Process." *BioScience* 37 (8): 550–56. https://doi.org/10.2307/1310665.

Friend, A. D., A. K. Stevens, R. G. Knox, and M. G. R. Cannell. 1997. "A Process-Based, Terrestrial Biosphere Model of Ecosystem Dynamics (Hybrid v3.0)." *Ecological Modelling* 95 (2): 249–87. https://doi.org/10.1016/S0304-3800(96)00034-8.

Friend, Andrew D., Wolfgang Lucht, Tim T. Rademacher, Rozenn Keribin, Richard Betts, Patricia Cadule, Philippe Ciais, et al. 2014. "Carbon Residence Time Dominates Uncertainty in Terrestrial Vegetation Responses to Future Climate and Atmospheric CO2." *Proceedings of the National Academy of Sciences* 111 (9): 3280–85. https://doi.org/10.1073/pnas.1222477110.

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): 1169–94.

Hammond, William M., A. Park Williams, John T. Abatzoglou, Henry D. Adams, Tamir Klein, Rosana López, Cuauhtémoc Sáenz-Romero, Henrik Hartmann, David D. Breshears, and Craig D. Allen. 2022. "Global Field Observations of Tree Die-off Reveal Hotter-Drought Fingerprint for Earth's Forests." *Nature Communications* 13 (1): 1761. https://doi.org/10.1038/s41467-022-29289-2.

Haralick, Robert M., K. Shanmugam, and Its'Hak Dinstein. 1973. "Textural Features for Image Classification." *IEEE Transactions on Systems, Man, and Cybernetics* SMC-3 (6): 610–21. https://doi.org/10.1109/TSMC.1973.4309314.

Haxeltine, Alex, and I. Colin Prentice. 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., C. Ryan, and M. Williams. 2015. "A Framework for Estimating Forest Disturbance Intensity from Successive Remotely Sensed Biomass Maps: Moving beyond Average Biomass Loss Estimates." *Carbon Balance and Management*. https://doi.org/10.1186/s13021-015-0039-0.

Hurtt, George C., PauL. R. Moorcroft, Stephen W. Pacala And, and Simon A. Levin. 1998. "Terrestrial Models and Global Change: Challenges for the Future." *Global Change Biology* 4 (5): 581–90. https://doi.org/10.1046/j.1365-2486.1998.t01-1-00203.x.

Jans, Luc, Lourens Poorter, Renaat S. A. R. van Rompaey, and Frans Bongers. 1993. "Gaps and Forest Zones in Tropical Moist Forest in Ivory Coast." *Biotropica* 25 (3): 258–69. https://doi.org/10.2307/2388784.

Jung, Martin, Christopher Schwalm, Mirco Migliavacca, Sophia Walther, Gustau Camps-Valls, Sujan Koirala, Peter Anthoni, et al. 2020. "Scaling Carbon Fluxes from Eddy Covariance Sites to Globe: Synthesis and Evaluation of the FLUXCOM Approach." *Biogeosciences* 17 (5): 1343–65. https://doi.org/10.5194/bg-17-1343-2020.

Köhler, Peter, and Andreas Huth. 1998. "The Effects of Tree Species Grouping in Tropical Rainforest Modelling: Simulations with the Individual-Based Model Formind." *Ecological Modelling* 109 (3): 301–21. https://doi.org/10.1016/S0304-3800(98)00066-0.

Lawton, Robert O., and Francis E. Putz. 1988. "Natural Disturbance and Gap-Phase Regeneration in a Wind-Exposed Tropical Cloud Forest." *Ecology* 69 (3): 764–77. https://doi.org/10.2307/1941025.

Le Toan, T., S. Quegan, M. W. J. Davidson, H. Balzter, P. Paillou, K. Papathanassiou, S. Plummer, et al. 2011. "The BIOMASS Mission: Mapping Global Forest Biomass to Better Understand the Terrestrial Carbon Cycle." *Remote Sensing of Environment*, DESDynI VEG-3D Special Issue, 115 (11): 2850–60. https://doi.org/10.1016/j.rse.2011.03.020.

Liu, Shuguang, Norman Bliss, Eric Sundquist, and Thomas G. Huntington. 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, Shuguang, Ben Bond-Lamberty, Jeffrey A. Hicke, Rodrigo Vargas, Shuqing Zhao, Jing Chen, Steven L. Edburg, et al. 2011. "Simulating the Impacts of Disturbances on Forest Carbon Cycling in North America: Processes, Data, Models, and Challenges." *Journal of Geophysical Research: Biogeosciences* 116 (G4). https://doi.org/10.1029/2010JG001585.

Malhi, Y., O. L. Phillips, O. L. Phillips, T. R. Baker, L. Arroyo, N. Higuchi, T. J. Killeen, 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, Nate G., Gerard Sapes, Alexandria Pivovaroff, Henry D. Adams, Craig D. Allen, William R. L. Anderegg, Matthias Arend, et al. 2022. "Mechanisms of Woody-Plant Mortality under Rising Drought, CO2 and Vapour Pressure Deficit." *Nature Reviews Earth & Environment*, March, 1–15. https://doi.org/10.1038/s43017-022-00272-1.

McGuire, A. D., J. M. Melillo, L. A. Joyce, D. W. Kicklighter, A. L. Grace, B. Moore III, and C. J. Vorosmarty. 1992. "Interactions between Carbon and Nitrogen Dynamics in Estimating Net Primary Productivity for Potential Vegetation in North America." *Global Biogeochemical Cycles* 6 (2): 101–24. https://doi.org/10.1029/92GB00219.

Moorcroft, P. R., G. C. Hurtt, and S. W. Pacala. 2001. "A Method for Scaling Vegetation Dynamics: The Ecosystem Demography Model (Ed)." *Ecological Monographs* 71 (4): 557–86. https://doi.org/10.1890/0012-9615(2001)071[0557:AMFSVD]2.0.CO;2.

Nash, J. E., and J. V. Sutcliffe. 1970. "River Flow Forecasting through Conceptual Models Part I — A Discussion of Principles." *Journal of Hydrology* 10 (3): 282–90. https://doi.org/10.1016/0022-1694(70)90255-6.

Nelson, Bruce W., Valerie Kapos, John B. Adams, Wilson J. Oliveira, and Oscar P. G. Braun. 1994. "Forest Disturbance by Large Blowdowns in the Brazilian Amazon." *Ecology* 75 (3): 853–58. https://doi.org/10.2307/1941742.

Norby, Richard J., Kiona Ogle, Peter S. Curtis, Franz-W. Badeck, Andreas Huth, George C.

Hurtt, Takashi Kohyama, and Josep Peñuelas. 2001. “Aboveground Growth and Competition in Forest Gap Models: An Analysis for Studies of Climatic Change.” *Climatic Change* 51 (3): 415–47. https://doi.org/10.1023/A:1012510619424.

Pan, Y., J. M. Chen, R. Birdsey, K. McCullough, L. He, and F. Deng. 2010. “Age Structure and Disturbance Legacy of North American Forests.” *Biogeoscience Discussions. 7: 979-1020.* 7: 979–1020. https://doi.org/10.5194/bgd-7-979-2010.

Parton, W. J., D. S. Schimel, C. V. Cole, and D. S. Ojima. 1987. “Analysis of Factors Controlling Soil Organic Matter Levels in Great Plains Grasslands.” *Soil Science Society of America Journal* 51 (5): 1173–79. https://doi.org/10.2136/sssaj1987.03615995005100050015x.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, et al. 2011. “Scikit-Learn: Machine Learning in Python.” *Journal of Machine Learning Research* 12 (85): 2825–30.

Potter, Christopher S., James T. Randerson, Christopher B. Field, Pamela A. Matson, Peter M. Vitousek, Harold A. Mooney, and Steven A. Klooster. 1993. “Terrestrial Ecosystem Production: A Process Model Based on Global Satellite and Surface Data.” *Global Biogeochemical Cycles* 7 (4): 811–41. https://doi.org/10.1029/93GB02725.

Raich, J. W., E. B. Rastetter, J. M. Melillo, D. W. Kicklighter, P. A. Steudler, B. J. Peterson, A. L. Grace, B. Moore III, and C. J. Vorosmarty. 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, Christian H., Veronika Gayler, Daniel Goll, Stefan Hagemann, Marvin Heidkamp, Julia E. M. S. Nabel, Thomas Raddatz, Erich Roeckner, Reiner Schnur, and Stiig Wilkenskjeld. 2021. “JSBACH 3 - The Land Component of the MPI Earth System Model: Documentation of Version 3.2,” February. https://doi.org/10.17617/2.3279802.

Rosen, Paul A., Scott Hensley, Scott Shaffer, Louise Veilleux, Manab Chakraborty, Tapan Misra, Rakesh Bhan, V. Raju Sagi, and R. Satish. 2015. “The NASA-ISRO SAR Mission - An International Space Partnership for Science and Societal Benefit.” In *2015 IEEE Radar Conference (RadarCon)*, 1610–13. https://doi.org/10.1109/RADAR.2015.7131255.

Runkle, James R. 2000. “Canopy Tree Turnover in Old-Growth Mesic Forests of Eastern North America.” *Ecology* 81 (2): 554–67. https://doi.org/10.2307/177448.

Running, Steven W., and Stith T. Gower. 1991. “FOREST-BGC, A General Model of Forest Ecosystem Processes for Regional Applications. II. Dynamic Carbon Allocation and Nitrogen Budgets1.” *Tree Physiology* 9 (1–2): 147–60. https://doi.org/10.1093/treephys/9.1-2.147.

Ryan, Casey M., Timothy Hill, Emily Woollen, Claire Ghee, Edward Mitchard, Gemma Cassells, John Grace, Iain H. Woodhouse, and Mathew Williams. 2012. “Quantifying Small-Scale Deforestation and Forest Degradation in African Woodlands Using Radar Imagery.” *Global Change Biology* 18 (1): 243–57. https://doi.org/10.1111/j.1365-2486.2011.02551.x.

Saatchi, Sassan S., Nancy L. Harris, Sandra Brown, Michael Lefsky, Edward T. A. Mitchard, William Salas, Brian R. Zutta, 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, Maurizio, Oliver Cartus, Nuno Carvalhais, Danaë M. A. Rozendaal, Valerio Avitabile, Arnan Araza, Sytze de Bruin, et al. 2021. “The Global Forest Above-Ground Biomass Pool for 2010 Estimated from High-Resolution Satellite Observations.” *Earth System Science Data* 13 (8): 3927–50. https://doi.org/10.5194/essd-13-3927-2021.

Senf, Cornelius, and Rupert Seidl. 2021a. “Storm and Fire Disturbances in Europe: Distribution and Trends.” *Global Change Biology* 27 (15): 3605–19.

https://doi.org/10.1111/gcb.15679.
———. 2021b. “Mapping the Forest Disturbance Regimes of Europe.” *Nature Sustainability* 4 (1): 63–70. https://doi.org/10.1038/s41893-020-00609-y.
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, Elena, Stephen W. Pacala, Sergey Malyshev, George C. Hurtt, P. C. D. Milly, John P. Caspersen, Lori T. Sentman, Justin P. Fisk, Christian Wirth, and Cyril Crevoisier. 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, Herman H., Rik Leemans, and Gordon B. Bonan, eds. 1992. *A Systems Analysis of the Global Boreal Forest*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511565489.
Spellerberg, Ian F., and Peter J. Fedor. 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): 177–79. https://doi.org/10.1046/j.1466-822X.2003.00015.x.
Stavros, E. Natasha, David Schimel, Ryan Pavlick, Shawn Serbin, Abigail Swann, Laura Duncanson, Joshua B. Fisher, et al. 2017. “ISS Observations Offer Insights into Plant Function.” *Nature Ecology & Evolution* 1 (7): 1–5. https://doi.org/10.1038/s41559-017-0194.
Tatarinov, Fyodor A., and Emil Cienciala. 2006. “Application of BIOME-BGC Model to Managed Forests: 1. Sensitivity Analysis.” *Forest Ecology and Management* 237 (1): 267–79. https://doi.org/10.1016/j.foreco.2006.09.085.
Thurner, Martin, Christian Beer, Nuno Carvalhais, Matthias Forkel, Maurizio Santoro, Markus Tum, and Christiane Schmullius. 2016. “Large-Scale Variation in Boreal and Temperate Forest Carbon Turnover Rate Related to Climate.” *Geophysical Research Letters* 43 (9): 4576–85. https://doi.org/10.1002/2016GL068794.
Turner, Monica G. 2010. “Disturbance and Landscape Dynamics in a Changing World.” *Ecology* 91 (10): 2833–49. https://doi.org/10.1890/10-0097.1.
Wang, Weile, Kazuhito Ichii, Hirofumi Hashimoto, Andrew R. Michaelis, Peter E. Thornton, Beverly E. Law, and Ramakrishna R. Nemani. 2009. “A Hierarchical Analysis of Terrestrial Ecosystem Model Biome-BGC: Equilibrium Analysis and Model Calibration.” *Ecological Modelling* 220 (17): 2009–23. https://doi.org/10.1016/j.ecolmodel.2009.04.051.
Williams, Mathew, Timothy C. Hill, and Casey M. Ryan. 2013. “Using Biomass Distributions to Determine Probability and Intensity of Tropical Forest Disturbance.” *Plant Ecology & Diversity* 6 (1): 87–99. https://doi.org/10.1080/17550874.2012.692404.
Williams, Mathew, Paul A. Schwarz, Beverly E. Law, James Irvine, and Meredith R. Kurpius. 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, Zhihong, Junan Li, Junjei Fei, Xun Mao, Peng Gao, and Yonglan Ding. 2005. “Study on the Adsorptive Catalytic Voltammetry of Emodin at a Carbon Paste Electrode.” *Analytical Letters* 38 (10): 1641–50. https://doi.org/10.1081/AL-200065808.
Yavitt, J., J. Battles, G. E. Lang, and D. H. Knight. 1995. “The Canopy Gap Regime in a Secondary Neotropical Forest in Panama.” https://doi.org/10.1017/S0266467400008853.