Tropical forest carbon balance: effects of field- and satellite-based mortality regimes on the dynamics and the spatial structure of Central Amazon forest biomass

OPEN ACCESS
IOP Publishing
Environmental Research Letters
Environ. Res. Lett. **9** (2014) 034010 (10pp)
doi:10.1088/1748-9326/9/3/034010

Tropical forest carbon balance: effects of field- and satellite-based mortality regimes on the dynamics and the spatial structure of Central Amazon forest biomass

Alan V Di Vittorio[1], Robinson I Negrón-Juárez[1,2], Niro Higuchi[3] and Jeffrey Q Chambers[1,4]

[1] Lawrence Berkeley National Laboratory, Earth Sciences Division, One Cyclotron Road, Mail Stop 84R0171, Berkeley, CA 94720, USA
[2] Department of Ecology and Evolutionary Biology, Tulane University, New Orleans, LA 70118, USA
[3] National Institute for Amazonian Research (INPA), Manaus, AM, Brazil
[4] Department of Geography, University of California, Berkeley, CA 94720, USA

E-mail: avdivittorio@lbl.gov

Received 17 December 2013, revised 7 February 2014
Accepted for publication 19 February 2014
Published 19 March 2014

Abstract

Debate continues over the adequacy of existing field plots to sufficiently capture Amazon forest dynamics to estimate regional forest carbon balance. Tree mortality dynamics are particularly uncertain due to the difficulty of observing large, infrequent disturbances. A recent paper (Chambers *et al* 2013 *Proc. Natl Acad. Sci.* **110** 3949–54) reported that Central Amazon plots missed 9–17% of tree mortality, and here we address 'why' by elucidating two distinct mortality components: (1) variation in annual landscape-scale average mortality and (2) the frequency distribution of the size of clustered mortality events. Using a stochastic-empirical tree growth model we show that a power law distribution of event size (based on merged plot and satellite data) is required to generate spatial clustering of mortality that is consistent with forest gap observations. We conclude that existing plots do not sufficiently capture losses because their placement, size, and longevity assume spatially random mortality, while mortality is actually distributed among differently sized events (clusters of dead trees) that determine the spatial structure of forest canopies.

Keywords: Amazon, biomass, forest, mortality, power law

1. Introduction

The world's forests have been identified as the primary terrestrial carbon sink, with tropical forests assimilating about half of the total forest carbon uptake (Pan *et al* 2011). However, these and similar estimates (Lewis *et al* 2009) of forest carbon balance have high levels of uncertainty, partially due to undercharacterized mortality regimes (Körner 2003, Fisher *et al* 2008). Tropical forest mortality regimes are commonly described by an average annual mortality rate, and less frequently include the distribution of disturbance size (Chambers *et al* 2004, Lloyd *et al* 2009, Chambers *et al* 2013). The mortality rate represents a bulk loss of trees or biomass while the disturbance size distribution represents a spatial clustering of this loss. The annual mortality rate has received more attention (Chambers *et al* 2004, Gloor *et al* 2009) than its spatial pattern due to infrequency of large disturbances (Chambers *et al* 2013), relatively short

duration of measurement (Gloor *et al* 2009), limited spatial extent of measurement (Nelson *et al* 1994, Espírito-Santo *et al* 2010), and relatively small plot size (Fisher *et al* 2008). While advances in remote sensing have increased scientific understanding of mortality regimes (Kellner *et al* 2011, Chambers *et al* 2013), questions remain regarding statistical distributions of disturbance size and their effects on spatial patterns of forest dynamics. Answering such questions is critical for understanding how the relationship between annual mortality and disturbance size distribution influences estimates of forest carbon balance.

Until recently, understanding of tropical forest mortality regimes has been based primarily on plot data, with a few exceptions that utilize large-area satellite and aircraft data, leaving a large knowledge gap for disturbances between ~0.07 and 35 ha in size (Nelson *et al* 1994, Espírito-Santo *et al* 2010, Kellner *et al* 2011, Morton *et al* 2011, Negrón-Juárez *et al* 2011, Chambers *et al* 2013). Fortunately, new remote sensing approaches fill this gap and allow more comprehensive analyses of spatial and temporal effects of disturbance on forest dynamics. Two key features have emerged from combining plot and remotely sensed data in the Central Amazon: (1) plot-based average annual mortality rates underestimate landscape-scale rates, and (2) plot-based frequency distributions of disturbance size generally miss succession-inducing gaps due to small size and limited census interval (Chambers *et al* 2013). Furthermore, several studies indicate that this frequency distribution follows a power law (Fisher *et al* 2008, Chambers *et al* 2009b, Lloyd *et al* 2009, Negrón-Juárez *et al* 2010, Kellner *et al* 2011, Chambers *et al* 2013), but the shape of this distribution has not been rigorously tested.

While advances have been made in mortality regime research, debate continues regarding the adequacy of existing field plot networks in the Amazon basin for estimating carbon balance (Fisher *et al* 2008, Chambers *et al* 2009b, Gloor *et al* 2009, Lloyd *et al* 2009, Chambers *et al* 2013). This debate bears directly on recent estimates of a tropical forest carbon sink based on field plot data (Lewis *et al* 2009, Pan *et al* 2011) and is highly contingent upon estimates of both annual mortality rate and disturbance size distribution because biomass accumulation is estimated as the difference between gains (from tree growth and ingrowth) and mortality losses. Furthermore, due to the relatively short period of time since plot establishment (e.g. ~11 years on average in the Amazon basin) and the low density of current field plots across vast domains (e.g. the Amazon basin) (Gloor *et al* 2009), forest growth simulations are critical for determining whether field measurements have adequate spatial and temporal coverage to capture long-term, regional trends that are characterized by slow gains and fast losses of biomass (Körner 2003).

Recent studies using individual-based forest growth models incorporating both annual mortality and disturbance size distributions have shown some limitations of field plots. Fisher *et al* (2008) simulated equilibrium landscape biomass to demonstrate that increasing the proportion of large-area disturbances can bias plot-based estimates toward biomass gains. A remotely sensed estimate of Central Amazon average annual mortality rate increased plot-only estimates by 9–17% because the largest recorded plot disturbances affected only 8 trees, while mortality events detected by satellite ranged from 8 to more than 7000 trees (Chambers *et al* 2013). While model estimates of forest biomass dynamics are highly sensitive to annual mortality rate, the increased mortality rate estimated by Chambers *et al* (2013) is well within the standard deviation of the plot-based annual mortality rate distribution (Chambers *et al* 2004). However, the impacts of different disturbance size distributions on spatial patterns of simulated biomass and stand age are relatively unknown. Nonetheless, such spatial patterns determine what information can be obtained from a sample of field plots.

In contrast to detailed research on mortality regimes, other studies have argued that plot-based estimates of annual mortality rate are sufficient for estimating Amazon forest carbon balance (Gloor *et al* 2009, Lloyd *et al* 2009). One such study, however, (Lloyd *et al* 2009) simply challenged a technical matter in Fisher *et al* (2008) that had no affect on carbon balance results because their methods were internally consistent (see Chambers *et al* 2013). Another study (Gloor *et al* 2009) conflated the annual mortality rate distribution with the frequency distribution of disturbance size. More specifically, Gloor *et al* (2009) accounted for only annual mortality rate, which requires the assumption that mortality is randomly distributed in space, and not the spatial clustering of mortality into individual, contiguous events. Furthermore, they did not effectively incorporate extra-plot mortality, and thus they found that their plot observations accurately detect biomass gains at the plot level. The dispute, however, is not whether the estimates show gains, but whether the measurements sufficiently capture landscape-level losses.

Clearly, the relationship between annual mortality rate and disturbance size distribution needs to be examined to develop appropriate landscape sampling designs for robust estimation of forest dynamics and carbon balance. Here we rigorously determine that the frequency distribution of disturbance size in the Central Amazon follows a power law, and use it to evaluate impacts of mortality regimes on spatial patterns of simulated biomass and stand age. We evaluate mortality regimes estimated from plot-only and merged plot and Landsat data. The higher, merged estimate of mortality significantly decreases the simulated live biomass, and including Landsat-detected disturbance events larger than eight trees shifts the landscape structure from a random pattern to a significantly clustered mosaic of stand age and biomass. We argue that the spatial clustering associated with the additional mortality due to large events further explains why existing plots do not adequately capture disturbance losses, leading to potential overestimates of Central Amazon forest carbon uptake.

2. Materials and methods

2.1. Study location and data

This study covers the Central Amazon region surrounding Manaus, Amazonas State, Brazil (3.11°S, 60.03°W). Plot data

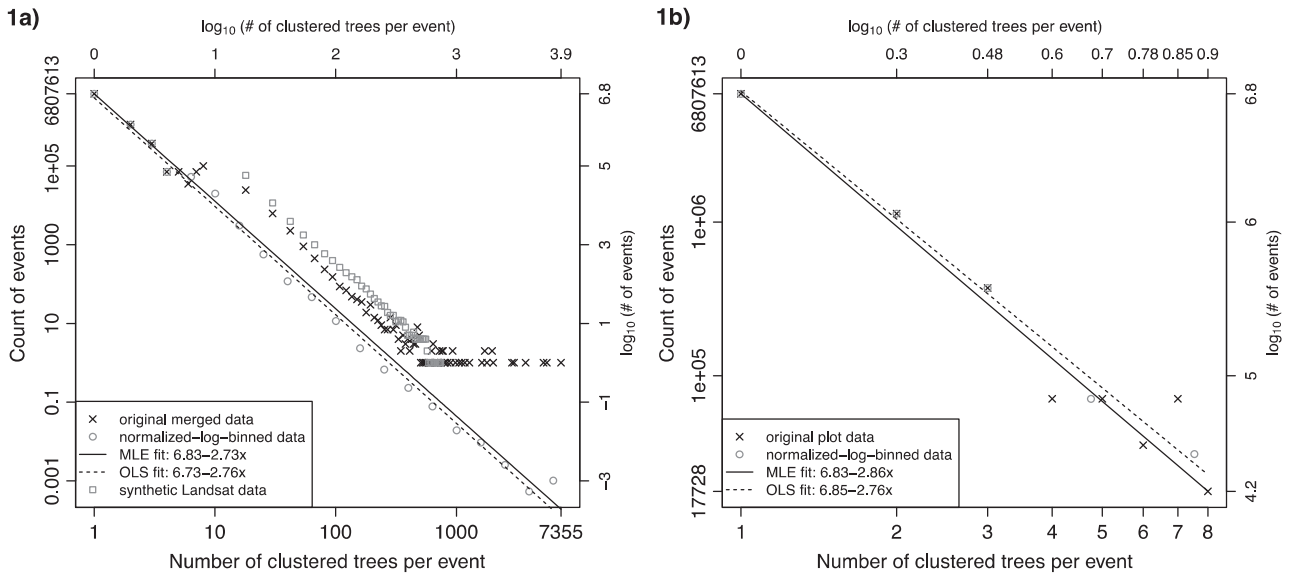

Figure 1. Mortality event distributions on $\log_{10} - \log_{10}$ axes. MLE = maximum likelihood estimator. OLS = ordinary least squares on normalized-log-binned data in $\log_{10} - \log_{10}$ space. (a) Merged data (original and normalized-log-binned) with the top two power law fits and a sample of synthetic Landsat data drawn from the MLE Landsat fit (power law exponent $\alpha = 3.5$). (b) Plot data (original and normalized-log-binned) with the top two power law fits. (a) Merged data. (b) Plot data.

Table 1. Mortality event distributions for the tropical tree ecosystem and community simulator (TRECOS).

Merged observations		Merged power law fit		Plot power law fit	
Event size class (# of trees)	Binned probability distribution function	Event size class (# of trees)	Binned probability distribution function	Event size class (# of trees)	Binned probability distribution function
1	$7.826\,453 \times 10^{-1}$	1	$7.900\,711\,3816 \times 10^{-1}$	1	$8.181\,1782 \times 10^{-1}$
2	$1.304\,409 \times 10^{-1}$	2	$1.190\,842\,0575 \times 10^{-1}$	2	$1.126\,8600 \times 10^{-1}$
3	$5.910\,607 \times 10^{-2}$	4	$6.707\,607\,3240 \times 10^{-2}$	3	$3.533\,8580 \times 10^{-2}$
7	$2.374\,431 \times 10^{-2}$	8	$1.710\,109\,0830 \times 10^{-2}$	4	$1.552\,1160 \times 10^{-2}$
18	$2.826\,444 \times 10^{-3}$	15	$4.577\,444\,2200 \times 10^{-3}$	5	$8.199\,0100 \times 10^{-3}$
35	$1.083\,556 \times 10^{-3}$	33	$1.664\,984\,0800 \times 10^{-3}$	6	$4.867\,4700 \times 10^{-3}$
84	$1.197\,948 \times 10^{-4}$	82	$3.382\,997\,2000 \times 10^{-4}$	7	$3.132\,1000 \times 10^{-3}$
207	$2.414\,290 \times 10^{-5}$	206	$6.912\,440\,0000 \times 10^{-5}$	8	$2.137\,8600 \times 10^{-3}$
589	$8.047\,633 \times 10^{-6}$	606	$1.644\,700\,0000 \times 10^{-5}$		
3128	$1.494\,560 \times 10^{-6}$	2763	$1.192\,600\,0000 \times 10^{-6}$		

have been collected in old-growth forest at permanent sites in a 50 km × 40 km area centered (2.5°S, 60°W) about 60 km north of Manaus (Chambers *et al* 2004). Standing dead trees accounted for 13.2% of total average annual plot-level tree mortality (1.02% of stems) while single and clustered wind-thrown trees comprised 50.6% and 36.2%, respectively (table 1 and figure 1). All estimates are based on trees having a diameter at breast height ≥ 10 cm. The largest mortality cluster in the observed plots contained eight trees.

Nine Landsat 5 Thematic Mapper scenes (Path 231, Row 062; $\sim 3.4 \times 10^4$ km^2 each; 30 m × 30 m resolution) were processed as five paired repeat observations of forest disturbance to obtain counts of wind-thrown tree clusters by size (1985–1986, 1987–1988, 1996–1997, 1997–1998, 2004–2005). These scenes were converted to reflectance values and masked to isolate 21 800 km^2 of old-growth forest area for analysis (aggregated across the five image pairs). Spectral mixture analysis was employed to calculate the fraction of shade-normalized non-photosynthetic vegetation (NPV) per pixel. Annual NPV change (ΔNPV) images were calculated and disturbed pixels were identified by ΔNPV ≥ 0.16 and then clustered into individual wind-throw events (e.g. Negrón-Juárez *et al* 2011). Aggregating the counts of wind-throw events across the five ΔNPV images generated a frequency distribution for mortality event sizes ranging from 8 to 7355 trees (Chambers *et al* 2013).

The plot and Landsat frequency distributions of wind mortality event size (as number of trees felled) were merged by extrapolating the plot data to the aggregate Landsat old-growth forest area and averaging the counts for the 8-tree event size (figure 1(a)). An evaluation of different merging methods on power law fits found that this method approximated the average

of all methods and that the spread in power law exponents had little effect on landscape dynamics and biomass. The merged frequency distribution spanned integer mortality event sizes, or classes, ranging from 1 to 7355 trees, and increased the estimate of total average annual tree mortality to 1.20% (Chambers *et al* 2013).

2.2. Mortality event distribution modeling

The three data sets (plot, Landsat, merged) were pre-processed to generate three data permutations for facilitating three different fit methods. (1) The original (unprocessed) data permutation was used by all methods. (2) The normalized-log-binned permutation summed counts in equal-sized $\log_{10}$ (class) bins and then divided these sums by the number of event class integers within each bin. (3) The normalized-binned permutation summed counts in equal-sized class bins and then divided these sums by the number of class integers within each bin. The binned data permutations reduce noise and generally eliminate zero-count classes (White *et al* 2008, Milojević 2010).

We fit power law ($y = c \cdot x^{-\alpha}$) and exponential ($y = c \cdot \exp(-\lambda \cdot x)$) functions to the plot, Landsat, and merged data sets, compared the fits using a likelihood ratio test, and tested the significance of a subset of fits using a bootstrap method. We used (1) a linear ordinary least squares (OLS) method on $\log_{10} - \log_{10}$ transformed values (White *et al* 2008, Milojević 2010) and (2) a discrete maximum likelihood estimator (MLE) method (Clauset *et al* 2009) to fit power law functions to the original and normalized-log-binned data permutations. We applied the MLE method using the observed minimum size class to ensure that all data were included in the fits. The exponential function was fit to the original and normalized-binned data permutations using (3) a linear OLS method on linear-$\log_e$ transformed values. We present analyses of the six most relevant fits for each method (tables 2 and 3).

For each data set (plot, Landsat, merged) we determined the best fit by comparing all possible pairs of fits using a likelihood ratio test (Clauset *et al* 2009). Based on the results of these comparisons we selected eight fits for which to calculate p-values signifying the probability that a random sample from the fitted distribution would have a worse fit to the original data than the fitted distribution. Each data fit was compared with 100 sample fits (p-value precision = 0.05) using the Kolmogorov–Smirnov D-statistic (KSD), which is the maximum distance between the cumulative distribution functions (CDF) of the data and the fitted function. If $p \leq 0.1$ we rejected the null hypothesis that the data follow the respectively fitted distribution function (Clauset *et al* 2009).

The random samples from the eight fits required different levels of processing to prepare the samples for fitting, depending on the data set, permutation, and fit method. The plot samples were used directly, the Landsat samples were binned to classes representing integer pixel estimates, and the merged samples were binned across the Landsat range with the 8-tree size class assigned the average of the binned and unbinned values. The eight selected fits generated the following sample fits: the MLE power law method was applied to original plot, Landsat, and merged samples, the OLS power law method was applied to normalized-log-binned plot, Landsat, and merged samples and the OLS exponential method was applied to original plot samples and normalized-binned plot samples.

2.3. Forest simulations

We used the tropical ecosystem and community simulator (TRECOS) (appendix; Chambers *et al* 2004 and Chambers *et al* 2013) to perform four 2000-year simulations of a Central Amazon forest landscape to evaluate the effects of two different mortality event distributions (table 1) and two different average annual mortality rates on landscape structure and biomass. The two average annual mortality rates were estimated from plot (1.02% of stems) and merged (1.20% of stems) data, respectively, and were modeled as normal distributions in $\log_{10}$ space using the merged estimate of inter-annual variability as the standard deviation (mean = 0.0086 (plot) and 0.0792 (merged) $\log_{10}$(% stems y^{-1}), SD = 0.073). The two event distributions were (1) the best fitted power law to the merged data and (2) the best fitted power law to the plot data. For input to TRECOS, we aggregated the merged event distribution to 10 bins and assigned each size class (1–8 trees) of the plot event distribution to its own bin.

We compared time series of mean landscape biomass and final spatial patterns among the simulations. For the time series we computed the average equilibrium biomass across simulation years 500–2000. We calculated a global clustering statistic (Moran's I) and empirical and fitted model semivariograms for each final-year map of biomass and the time since last succession-inducing disturbance (>8 trees, t_d) (R-project packages gstats and spdep; www.r-project.org/). Moran's I calculation covered the entire landscape and used a spatial weighting function ($1/(2 \cdot \text{distance})$) to emphasize clusters on the order of one hectare. Semivariogram calculations extended to half the distance across the landscape (500 m) to minimize bias due to decreasing sample sizes with increasing distance. Final-year biomass exhibited a log-normal distribution in space and t_d exhibited a fat-tail distribution, so we performed Monte Carlo significance tests for Moran's I (100 simulations each). We also compared histograms and mean/median statistics of these maps.

3. Results

The OLS power law fits to normalized-log-binned data are similar to the MLE fits, except for the larger magnitude MLE exponents for Landsat data sets (tables 2 and 3). The OLS power law fit to the original plot data is also similar to the observed value MLE fit. The OLS exponential fits to the plot data permutations are somewhat plausible, while the exponential fits to other data sets are not (table 3). All regressions are significant to the 95% confidence level. It is apparent that the OLS fits are valid only on appropriately binned data or 'full' plot data sets with values for all integer classes. Henceforth, references to OLS fits refer to those performed on binned data unless otherwise noted.

Table 2. Maximum likelihood estimator (MLE) power law probability distribution function fits ($y = c \cdot x^{-\alpha}$).

Data configuration	Minimum size class (MSC)	Power law exponent (α)	Normalization factor (c)	Goodness of fit (R^2)	Sample size (n)
Plot data[a]	1	2.86	$8.114\,722 \times 10^{-1}$	0.9987	8580 429
Landsat data[a,b]	8	3.50	$3.875\,981 \times 10^{2}$	0.9935	217 944
Merged data[a]	1	2.73	$7.900\,710 \times 10^{-1}$	0.9994	8698 210
Plot data (NLB[c])	1	3.01	$8.332\,725 \times 10^{-1}$	0.9970	8416 444
Landsat data (NLB)	8	3.50	$3.875\,981 \times 10^{2}$	0.9998	48 703
Merged data (NLB)	1	2.97	$8.277\,390 \times 10^{-1}$	0.9981	8440 758

[a] The ratio tests rank the MLE fits on original data above all others.
[b] Null hypothesis of power law distribution rejected ($p < 0.1$). P-value tests were not performed for any of the NLB fits.
[c] NLB = Normalized-log-binned. The plot and merged results are for a start bin center of 3 and are nearly identical to those for a start bin center of 4. The Landsat data have a start bin center of 8.

Table 3. Ordinary least squares (OLS) power law ($y = c \cdot x^{-\alpha}$) and exponential ($y = c \cdot \exp(-\lambda \cdot x)$) probability distribution function fits[a].

Data and function configuration	Slope ($-\alpha$ or $-\lambda$)	Intercept	Normalization factor (c)	Coefficient of determination for linear fit (R^2)	Sample size (n)
Plot data					
Power law	−2.76	6.81	$7.947\,239 \times 10^{-1}$	0.9505	8
Exponential[b]	−0.75	15.39	$1.106\,441 \times 10^{0}$	0.8410	8
Landsat data					
Power law	−1.79	5.53	$3.859\,451 \times 10^{0}$	0.7917	71
Exponential	−0.001	2.49	$1.007\,528 \times 10^{-3}$	0.1347	71
Merged data					
Power law	−2.00	6.11	$6.072\,334 \times 10^{-1}$	0.9108	78
Exponential	−0.001	3.74	$1.000\,500 \times 10^{-3}$	0.1515	78
Plot data (Binned[c])					
Power law (s.b. = 3)	−2.76	6.85	$7.950\,641 \times 10^{-1}$	0.9932	5
Exponential[b] (s.b. = 3)	−0.87	16.32	$1.386\,911 \times 10^{0}$	0.9813	4
Landsat data (Binned)					
Power law[b] (s.b. = 8)	−2.71	6.67	$5.408\,683 \times 10^{1}$	0.99	15
Exponential (s.b. = 8)	−0.001	−0.51	$1.007\,528 \times 10^{-3}$	0.1103	68
Merged data (Binned)					
Power law (s.b. = 4)	−2.76	6.73	$7.945\,536 \times 10^{-1}$	0.9940	20
Exponential	−0.001	−0.42	$1.007\,528 \times 10^{-3}$	0.1062	68

[a] OLS was performed on $\log_{10} - \log_{10}$ transformed data for power law fits, and on linear-$\log_e$ transformed data for exponential fits.
[b] Null hypothesis of power law or exponential distribution rejected ($p < 0.1$). For the exponential distribution, P-value tests were performed only for plot data fits (binned and non-binned). For the power law distribution, P-value tests were performed only for the binned data.
[c] The data are normalized-log-binned for the power law fits and normalized-binned for the exponential fits. The plot and merged results are for a start bin (s.b.) center of 3 unless otherwise noted, and in general are nearly identical to those for a start bin center of 4. The Landsat data have a start bin center of 8. The binned data are transformed as noted.

The likelihood ratio tests rank the MLE power law fits to original data above all other fits, with plot and merged data having the best fits (table 2, figure 1). The highest ranked OLS power law fits to the plot and merged datasets are effectively identical to the MLE fits. All OLS exponential fits are ranked below all power law fits. The p-values indicate that only plot and merged data are likely to follow their respectively fitted power law distributions (tables 2 and 3). Thus, we reject the hypotheses that plot data follow an exponential distribution and that the given Landsat data follow a power law or an exponential distribution.

The mortality event distribution (table 1) and the average annual mortality rate each affect the simulated landscape differently. Increasing average annual mortality decreases landscape-level biomass (figure 2(a)) but has little effect on t_d and spatial pattern. Using the merged mortality event distribution, which includes large-area events, causes local spatial autocorrelation of biomass and t_d (figures 2(b)–(c) and 3) but has little effect on mean (figure 2(a)) and median values of these variables. The semivariogram results (figure 3) are supported by significantly positive, although small, Moran's I values (figures 2(b) and (c)). In contrast, using the plot-based mortality event distribution generates relatively flat semivariograms and insignificant Moran's I values, indicating that spatially random outputs are generated when only plot data are used as input.

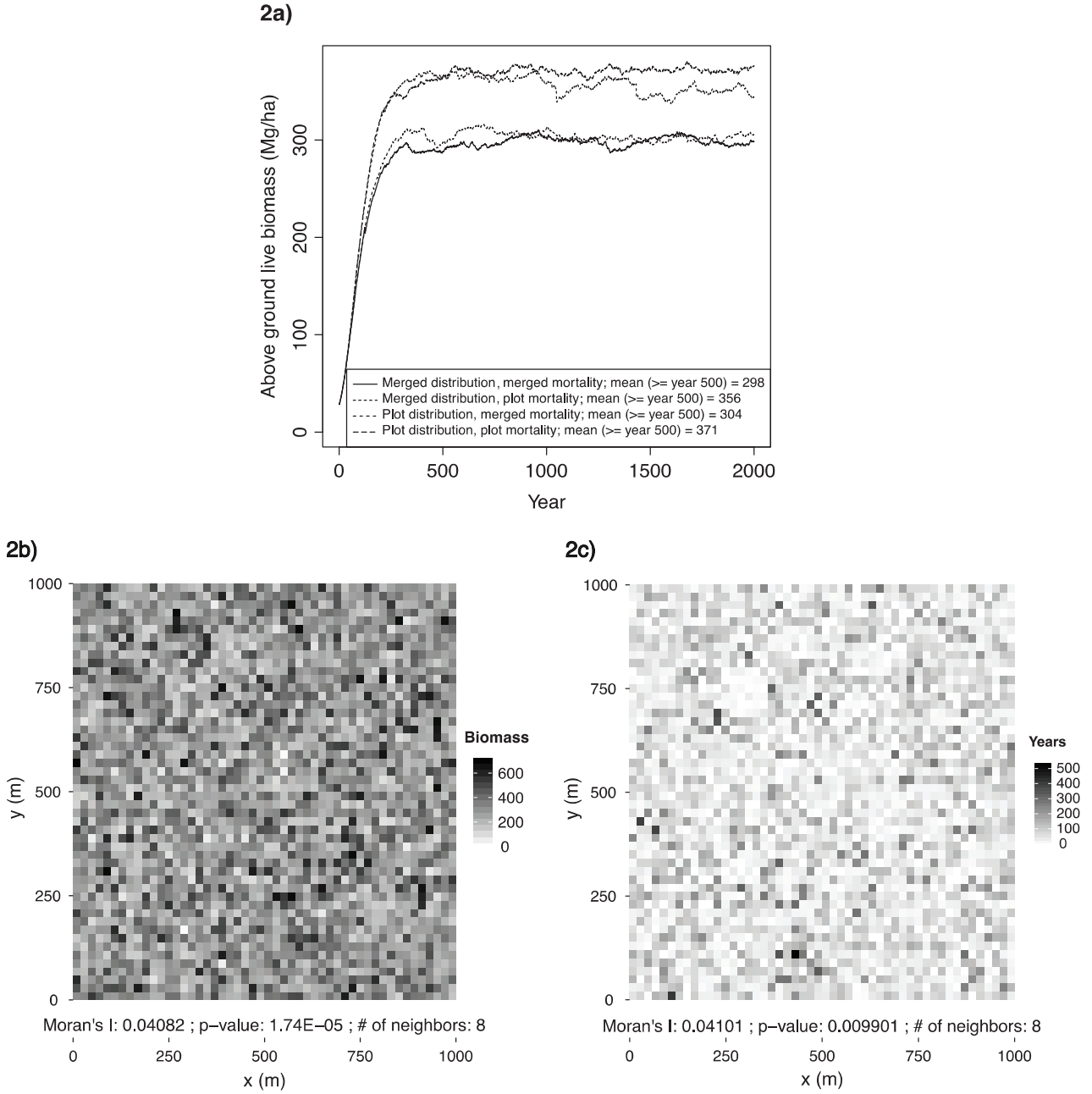

Figure 2. Temporal and spatial results of a forest growth model. (a) Time series of mean above ground biomass for a 100 ha landscape. Higher annual stem mortality associated with the merged data reduces biomass, regardless of the event size distribution. (b) Biomass map of 2500 forest stand cells (400 m^2 each) for year 2000 of the merged distribution (table 1), merged average annual stem mortality (1.20%) simulation. Only 0.92% of the cells have biomass values >700 Mg ha^{-1}. The corresponding map using plot average annual stem mortality (1.02%) is also significantly clustered ($I = 0.001\,559$, p-value $= 0.009\,901$). (c) Map of years since last disturbance >8 dead trees for year 2000 of the merged distribution (table 1), merged average annual stem mortality (1.20%) simulation. Significantly positive Moran's I value indicates spatial clustering. The corresponding map using plot average annual stem mortality (1.02%) is also significantly clustered ($I = 0.000\,831$, p-value $= 0.0198$). (a) Mean above ground live biomass. (b) Above ground live biomass (Mg ha^{-1}). (c) Years since last event.

4. Discussion

The results overwhelmingly show that Central Amazon mortality events can be robustly modeled by a power law distribution function, and that an exponential function is not adequate even for plot data. The merged data set is the most comprehensive mortality data available for this region and the power law exponent of its best fit (−2.73) is consistent with previous estimates of gap area frequency in tropical forests (Fisher *et al* 2008, Kellner and Asner 2009, Lloyd *et al* 2009). While research consistently shows that mortality events and gap areas follow a power law distribution, the mechanisms behind this distribution are still unclear.

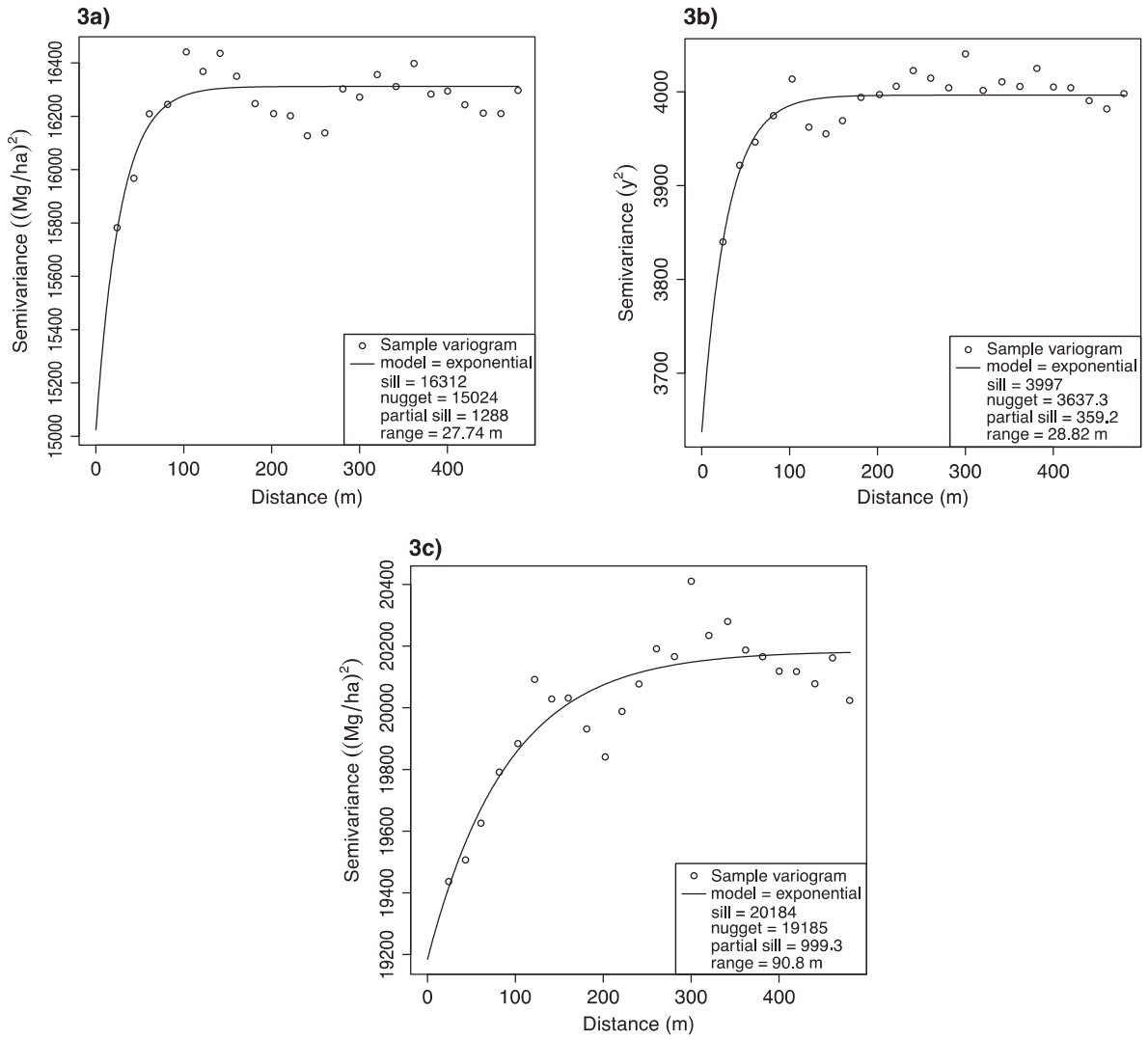

Figure 3. Semivariograms of simulation year 2000 above ground biomass and disturbance interval maps. There are 2500 forest stand cells (400 m^2 each). The model semivariogram range estimates the radius at which 95% of the partial sill has been reached. (a) The merged event distribution (table 1) with merged average annual stem mortality (1.20%) generates spatial autocorrelation of biomass within one and a half forest cells (range $= 27.74$ m^2). (b) The merged event distribution with merged average annual stem mortality generates spatial autocorrelation of disturbance interval within one and a half forest cells (range $= 28.82$ m^2), which corresponds with the respective biomass autocorrelation (a). (c) The merged event distribution with plot average annual stem mortality (1.02%) generates spatial autocorrelation of biomass within four and a half forest cells (range $= 90.8$ m^2). The disturbance interval empirical semivariogram corresponding to (c) (not shown) continually increases for distances >200 m. (a) Biomass: Merged distribution, merged mortality. (b) Event interval: Merged distribution, merged mortality. (c) Biomass: Merged distribution, plot mortality.

However, insightful comparisons of gap frequency distributions are difficult to make because of different methods and gap definitions. Most tropical forest mortality distributions are based on gap area frequency (Fisher *et al* 2008, Kellner and Asner 2009, Lloyd *et al* 2009), but gap area is an indirect measure of mortality due to spatial and temporal variation in tree density (Chapman *et al* 1997), and pixel footprints pose challenges for remotely measuring gap area because entire pixels are rarely devoid of trees (Nelson *et al* 1994). We use instead the dead tree count to directly characterize mortality within one year of gap formation, and our Landsat method generates a good relationship between ΔNPV and number of dead trees within each pixel, enabling repeat measurements of sub-pixel mortality. It is also very well suited for merging with plot measurements to create a contiguous mortality event size distribution ranging from 1 to over 7000 trees, with the potential to capture events larger than 2000 ha (on the order of 400 000 trees) with additional data (Espírito-Santo *et al* 2010).

High-resolution LiDAR (Light Detection And Ranging) captures forest structure well (Kellner and Asner 2009), but its ability to directly measure large disturbance-induced gaps is limited by high data volume to small areas and low temporal frequency (Kellner *et al* 2011). Nonetheless, one-time measurements of canopy gaps having vegetation height <1 m are linked to gap formation unless edaphic or other non-disturbance factors limit these areas to short vegetation. LiDAR-based analysis of gaps with vegetation height <1 m in four 1 km^2 tropical landscapes estimated power law gap

frequency distributions with exponents ranging from −2.00 to −2.33 (Kellner and Asner 2009). The maximum gap size was on the order of 10 000 m^2 (1 ha) with most gaps ≤100 m^2 (0.01 ha), which corresponds well with the maximum mortality event size of our plot data (8 trees where average tree density is 605 trees ha^{-1}). The best fitted power law exponent for our plot data (−2.86) is more negative than the (Kellner and Asner 2009) exponents, but these two results are remarkably similar considering that ours is based on recently fallen trees and the other is based on contiguous vegetation height.

Other estimates of power law exponents for gap frequency distributions are more or less comparable to ours even though methods differ. Some exponents for gaps smaller than 1000 m^2 (0.1 ha) have been estimated by OLS for logarithmically binned data rather than by MLE on the original data to provide a binned input distribution for an empirical gap dynamics model (Fisher *et al* 2008, Negrón-Juárez *et al* 2010). These OLS estimates ranged from −1.1 to −1.6 and are mathematically equivalent to the power law exponent plus one (White *et al* 2008, Chambers *et al* 2013). When transformed to the power law exponent (−2.1 to −2.6) these estimates are comparable to MLE fits of the same source data (−1.9 to −2.7; Lloyd *et al* 2009), and also to our best fits (table 2) for plot and merged data. However, Lloyd *et al* (2009) estimated an exponent of −3.1 for remotely sensed gaps ranging from 32 to over 1 700 ha (320 000 to 17 000 000 m^2; Nelson *et al* 1994), which generally are much larger than our largest mortality event of about 35 ha (350 000 m^2). Even with discrepancies among methods and gap definitions, the consistency in exponents among these studies strongly supports our result that tropical forest mortality event size distributions follow a power law, and that these distributions extend to large mortality events that have not been adequately captured by existing field plots. Thus, the spatial and temporal domains sampled by existing field plots are generally too small to accurately estimate the mortality biomass flux.

This study demonstrates that both MLE and OLS methods for fitting power law functions can be reliable if the data are appropriate to the method used. The MLE algorithm finds the best power law fit to a complete data set i.e. there are observations of each discrete event size, as shown by the ratio and p-values tests (tables 2 and 3). The Landsat data, however, are inherently binned because each observed event size is based on an average mortality rate associated with an integral number of pixels. For example, an isolated mortality pixel might contain 4–19 dead trees (Negrón-Juárez *et al* 2011), but it is counted as an 8-tree cluster. This binning renders the Landsat data inappropriate for the MLE method, and as a result the MLE fit to the original Landsat data is not statistically robust based the p-value metrics, although this lack of significance could be an artifact of the noisy tail.

The OLS normalized-log-binned method, on the other hand, can estimate good power law fits for the plot and merged data sets (table 3 and figure 1) and at first glance appears to be more reliable than the MLE method for the Landsat data. The binning process helps account for non-sampled event sizes, smoothens out the noisy tail (figure 1), and redistributes pre-binned counts such as occur in the Landsat data. The OLS power law fits to the plot data are good, but MLE is superior for such a complete sample. The close match of the MLE and OLS power law estimates for the merged data (1% difference) indicates that these are both valid, and likely the best, estimates of mortality event distribution for this region. The OLS Landsat power law exponents are more similar to the valid merged data exponents than the MLE Landsat exponents, suggesting that in this case OLS outperforms MLE. But a binned sample from the MLE Landsat fitted distribution, which has a much steeper slope than the OLS fit (tables 2 and 3), more closely matches the observed data, albeit without the noisy tail (figure 1). And even though all p-value tests reject the power law hypotheses for Landsat data, the ratio test ranks the MLE Landsat fit higher than the OLS fit (tables 2 and 3). Discarding the noisy counts for event sizes greater than 458 trees (∼2.88 ha or 28 800 m^2) might generate a more reliable fit to the Landsat data, but it would severely restrict the range of event sizes represented by the fitted function. Without a larger data set it is difficult to determine how well the normalized-log-binned data represent the noisy tail and thus whether the Landsat data is adequately represented by the merged data fit or should be represented separately with a steeper slope. However, both MLE and OLS fits to the merged data do incorporate all observations consistently (figure 1).

These robust mortality event distributions enable us to simulate forest landscapes and to show that a more complete event distribution generates spatial patterns in biomass and stand age, in contrast to simulations using a plot-estimated distribution (figures 2 and 3). This spatial difference is augmented by the expected result that a higher average annual mortality rate, determined from the merged data set, reduces above ground biomass in relation to that obtained from the plot-based rate (figure 2(a)). Using the higher mortality rate in simulations slightly underestimates landscape biomass averages in relation to wider area plot observations (∼319 Mg ha^{-1}; Chambers *et al* 2004), in contrast to overestimation when using the lower average mortality rate (figure 2). The higher average mortality rate also reduces the range of spatial autocorrelation due to the higher frequency of small events over large events in the power law distribution (figure 3). It is therefore apparent that annual rates and event distributions are both key characteristics for understanding tropical forest dynamics.

When using the merged data estimates for average morality rate and event distribution, the simulated spatial patterns are consistent with observations of large gaps, but the forest stand cell size used in our analyses is too large to adequately capture all spatial structure of the forest. The estimated range of spatial autocorrelation for both biomass and t_d is ∼28 m (figures 3(a)–(b)), which corresponds to a circular area of 616 m^2. As mentioned above, the largest gap with vegetation height less than 1 m detected by LiDAR in four tropical landscapes was 10 000 m^2 (1 ha), with most gaps ≤ 100 m^2 (0.01 ha) (Kellner and Asner 2009). Thus, the TRECOS simulations using merged data adequately capture the observed, large-area spatial coherence in tropical forest. Additionally, Kellner *et al* (2011) estimated ranges of spatial autocorrelation for canopy height (8.3–21.1 m) and changes in canopy height (2.5–3.7 m) for forests on five different substrate ages in

Hawaii. Unfortunately these canopy range estimates, and the majority of gaps, are smaller than what TRECOS can resolve because the stand cells are 20 m × 20 m. Nonetheless, the merged disturbance size distribution is required to simulate observed spatial patterns, demonstrating that mortality is not randomly distributed across the landscape

5. Conclusion

Our results show that existing plot data alone are not sufficient for characterizing tropical forest mortality processes and their impacts on landscape-level biomass density and also on spatial patterns of succession-inducing disturbance interval and biomass density. A power law distribution of mortality event size based on merged plot and satellite data is required in addition to the corresponding annual mortality rate to shift simulated forest structure from a random pattern to a significantly clustered mosaic of stand age and biomass. Thus, we conclude that the spatial clustering associated with mortality due to events larger than eight trees further explains why existing plots do not adequately capture large disturbance losses that influence succession, canopy structure, growth rates, and species composition (Chambers *et al* 2009a) Due to this underestimation of mortality, current plot-based analyses have a tendency to overestimate forest carbon uptake. It is apparent that a more comprehensive sampling scheme that includes large-area data (e.g., large plots and remote sensing) and robustly characterizes disturbance size distribution is required to understand tropical forest dynamics and its impact on carbon balance.

Acknowledgments

This project was funded by the Director, Office of Science, Office of Biological and Environmental Research, of the US Department of Energy under contract No. DE-AC02-05CH11231 as part of the Integrated Assessment Research Program.

Appendix. Stand growth model

The tropical ecosystem and community simulator (TRECOS) (Chambers *et al* 2004, 2013) is a stochastic-empirical model that simulates growth and size distribution of individual trees grouped into 400 m^2 stands. It also simulates dead tree decomposition, tree size distribution (i.e. stem diameter), mortality, and recruitment. It has been developed and parameterized using field data from the plots used in this study, and it requires distributions of wind mortality event size (number of trees per event), overall annual stem mortality rate, diameter growth rate (mean $= -0.976 \log_{10}(\text{cm y}^{-1})$, SD $= 0.431$), stem density (mean $= 24.22$ stems/400 m^2, SD $= 5.11$), and wood density (mean $= 0.55$ (early succession) and $0.7\ \text{g cm}^{-3}$, SD $= 0.15$) as inputs. The standing dead stem mortality rate is applied randomly in space before the wind mortality event size distribution is used to allocate the remaining annual mortality. Biomass is calculated from allometric equations based on stem diameter. We did not simulate CO_2 fertilization effects in this study. Outputs include time series of live and dead biomass stocks and reproduction and mortality rates averaged over a 1 km^2 landscape and also for selected sample plots. TRECOS also outputs stand-resolution maps of the final biomass state and the time since last succession-inducing disturbance (>8 trees; t_d).

References

Chambers J Q, Higuchi N, Teixeira L M, Santos J d, Laurance S G and Trumbore S E 2004 Response of tree biomass and wood litter to disturbance in a Central Amazon forest *Oecologia* **141** 596–611

Chambers J Q, Negrón-Juárez R I, Hurtt G C, Marra D M and Higuchi N 2009b Lack of intermediate-scale disturbance data prevents robust extrapolation of plot-level tree mortality rates for old-growth tropical forests *Ecol. Lett.* **12** E22–5

Chambers J Q, Negrón-Juárez R I, Marra D M, Di Vittorio A, Tews J, Roberts D, Ribeiro G H P M, Trumbore S E and Higuchi N 2013 The steady-state mosaic of disturbance and succession across an old-growth Central Amazon forest landscape *Proc. Natl Acad. Sci.* **110** 3949–54

Chambers J, Robertson A, Carneiro V, Lima A, Smith M-L, Plourde L and Higuchi N 2009a Hyperspectral remote detection of niche partitioning among canopy trees driven by blowdown gap disturbances in the Central Amazon *Oecologia* **160** 107–17

Chapman C A, Chapman L J, Wrangham R, Isabirye-Basuta G and Ben-David K 1997 Spatial and temporal variability in the structure of a tropical forest *Afr. J. Ecol.* **35** 287–302

Clauset A C, Shalizi C R and Newman M E J 2009 Power-law distributions in empirical data *Soc. Ind. Appl. Math. Rev.* **51** 661–703

Espírito-Santo F D B, Keller M, Braswell B, Nelson B W, Frolking S and Vicente G 2010 Storm intensity and old-growth forest disturbances in the Amazon region *Geophys. Res. Lett.* **37** L11403

Fisher J I, Hurtt G C, Thomas R Q and Chambers J Q 2008 Clustered disturbances lead to bias in large-scale estimates based on forest sample plots *Ecol. Lett.* **11** 554–63

Gloor M *et al* 2009 Does the disturbance hypothesis explain the biomass increase in basin-wide Amazon forest plot data? *Glob. Change Biol.* **15** 2418–30

Kellner J R and Asner G P 2009 Convergent structural responses of tropical forests to diverse disturbance regimes *Ecol. Lett.* **12** 887–97

Kellner J, Asner G, Vitousek P, Tweiten M, Hotchkiss S and Chadwick O 2011 Dependence of forest structure and dynamics on substrate age and ecosystem development *Ecosystems* **14** 1156–67

Körner C 2003 Slow in, rapid out–carbon flux studies and Kyoto targets *Science* **300** 1242–3

Lewis S L *et al* 2009 Increasing carbon storage in intact African tropical forests *Nature* **457** 1003–6

Lloyd J, Gloor E U and Lewis S L 2009 Are the dynamics of tropical forests dominated by large and rare disturbance events? *Ecol. Lett.* **12** E19–21

Milojević S 2010 Power law distributions in information science: making the case for logarithmic binning *J. Am. Soc. Inform. Sci. Technol.* **61** 2417–25

Morton D C, DeFries R S, Nagol J, Souza C M Jr, Kasischke E S, Hurtt G C and Dubayah R 2011 Mapping canopy damage from understory fires in Amazon forests using annual time series of Landsat and MODIS data *Remote Sens. Environ.* **115** 1706–20

Negrón-Juárez R I, Chambers J Q, Guimaraes G, Zeng H, Raupp C F M, Marra D M, Ribeiro G H P M, Saatchi S S, Nelson B W and Higuchi N 2010 Widespread Amazon forest tree mortality from a single cross-basin squall line event *Geophys. Res. Lett.* **37** L16701

Negrón-Juárez R I, Chambers J Q, Marra D M, Ribeiro G H P M, Rifai S W, Higuchi N and Roberts D 2011 Detection of subpixel treefall gaps with Landsat imagery in Central Amazon forests *Remote Sens. Environ.* **115** 3322–8

Nelson B W, Kapos V, Adams J B, Wilson J O and Braun O P G 1994 Forest disturbance by large blowdowns in the Brazilian Amazon *Ecology* **75** 853–8

Pan Y, Chen J M, Birdsey R, McCullough K, He L and Deng F 2011 Age structure and disturbance legacy of North American forests *Biogeosciences* **8** 715–32

White E P, Enquist B J and Green J L 2008 On estimating the exponent of power-law frequency distributions *Ecology* **89** 905–12