Does the disturbance hypothesis explain the biomass increase in basin-wide Amazon forest plot data?

M. GLOOR*, O. L. PHILLIPS*, J. J. LLOYD*, S. L. LEWIS*, Y. MALHI†, T. R. BAKER*, G. LÓPEZ-GONZALEZ*, J. PEACOCK*, S. ALMEIDA‡, A. C. ALVES DE OLIVEIRA§, E. ALVAREZ¶, I. AMARAL§, L. ARROYO‖, G. AYMARD**, O. BANKI††, L. BLANC‡‡, D. BONAL‡‡, P. BRANDO§§, K.-J. CHAO¶¶, J. CHAVE‖‖, N. DÁVILA***, T. ERWIN†††, J. SILVA‡‡‡, A. DI FIORE§§§, T. R. FELDPAUSCH*, A. FREITAS¶¶, R. HERRERA‖‖‖, N. HIGUCHI****, E. HONORIO††††, E. JIMÉNEZ‡‡‡‡, T. KILLEEN§§§§, W. LAURANCE¶¶¶, C. MENDOZA‖‖‖‖, A. MONTEAGUDO*****, A. ANDRADE†††††, D. NEILL‡‡‡‡‡, D. NEPSTAD§§§§§, P. NÚÑEZ VARGAS¶¶¶¶, M. C. PEÑUELA‖‖‖‖‖, A. PEÑA CRUZ*****, A. PRIETO******, N. PITMAN††††††, C. QUESADA‡‡‡‡‡‡, R. SALOMÃO§§§§§§, MARCOS SILVEIRA¶¶¶¶¶¶, M. SCHWARZ‖‖‖‖‖‖, J. STROPP††, F. RAMÍREZ*******, H. RAMÍREZ†††††††, A. RUDAS******, H. ter STEEGE‡‡‡‡‡‡‡, N. SILVA§§§§§§, A. TORRES¶¶¶¶¶¶¶, J. TERBORGH‖‖‖‖‖‖‖, R. VÁSQUEZ***** and G. van der HEIJDEN********

*School of Geography, University of Leeds, Leeds, UK, †University of Oxford, Oxford, UK, ‡Paraense Emilio Goeldi, Belem, Brazil, §Team-INPA, Manaus, Brazil, ¶University of Colombia, Medellín, Colombia, ‖Missouri Botanical Garden and Museo de Historia Natural Noel Kempff Mercado, Santa Cruz, Bolivia, **UNELLEZ-Guanare, Venezuela, ††Utrecht University, the Netherlands, ‡‡UMR, Kourou, French Guiana, §§IPAM, Brazil, ¶¶Institute of Ecology and Evolutionary Biology, Taiwan, ‖‖Laboratoire de Evolution et Diversité Biologique, France, ***Universidad Nacional de la Amazonia Peruana, Peru, †††Smithsonian Institution, Washington, DC, USA, ‡‡‡Universidad de San Abad del Cusco, Peru, §§§New York University, New York, USA, ¶¶¶Emilio Goeldi Museum, Belem, Brazil, ‖‖‖IVIC, Caracas, Venezuela, ****Instituto National de Pesquisas de Amazonia, Manaus, Brazil, ††††IIAP, Iquitos, Peru, ‡‡‡‡Universidad Nacional de Colombia, Leticia, Colombia, §§§§Conservation International, Washington, DC, USA, ¶¶¶¶Smithsonian Tropical Research Institute, Ancon, Panama, ‖‖‖‖Fomabo, Manejo Forestal en las Tierras Tropicales de Bolivia, Sacta, Bolivia, *****Proyecto Flora del Perú, Jardin Botanico de Missouri, Oxapampa, Peru, †††††INPA, Manaus, Brazil, ‡‡‡‡‡Fundación Jatun Sacha and Missouri Botanical Garden, Quito, Ecuador, §§§§§Woods Hole Research Center, MA, USA, ¶¶¶¶¶Herbario Vargas, Universidad San Antonio Abad del Cusco, Cusco, Peru, ‖‖‖‖‖Universidad Nacional de Colombia, Leticia, Amazonas, ******Instituto de Ciencias Naturales, Universidad Nacional de Colombia, Bogotá, Colombia, ††††††Center for Tropical Conservation, Duke University, Durham, ‡‡‡‡‡‡School of Geography, University of Leeds, Leeds, UK, §§§§§§Museu Paraense Emilio Goeldi, Belem, Pará, ¶¶¶¶¶¶Universidade Federal do Acre, Acre, Brazil, ‖‖‖‖‖‖Jena, Germany, *******Universidad Nacional de la Amazonia Peruana, Iquitos, Peru, †††††††University of the Andes, Merida, Venezuela, ‡‡‡‡‡‡‡NHN Utrecht, the Netherlands, §§§§§§§CIFOR, Tapajos, Brasil & EMBRAPA Amazonia Oriental, Belém, Brazil, ¶¶¶¶¶¶¶Universidad de Los Andes, Merida, Venezuela, ‖‖‖‖‖‖‖Duke University, NC, USA, ********Leeds, UK

Abstract

Positive aboveground biomass trends have been reported from old-growth forests across the Amazon basin and hypothesized to reflect a large-scale response to exterior forcing. The result could, however, be an artefact due to a sampling bias induced by the nature of forest growth dynamics. Here, we characterize statistically the disturbance process in Amazon old-growth forests as recorded in 135 forest plots of the RAINFOR network up to 2006, and other independent research programmes, and explore the consequences of sampling artefacts using a data-based stochastic simulator. Over the observed range of annual aboveground biomass losses, standard statistical tests show that the distribution of biomass losses through mortality follow an exponential or near-identical Weibull probability distribution and not a power law as assumed by others. The simulator was parameterized using both an exponential disturbance probability distribution as well as a mixed exponential–power law distribution to account for potential large-scale blow-

down events. In both cases, sampling biases turn out to be too small to explain the gains detected by the extended RAINFOR plot network. This result lends further support to the notion that currently observed biomass gains for intact forests across the Amazon are actually occurring over large scales at the current time, presumably as a response to climate change.

Introduction

Humans are in the process of significantly altering the global atmospheric environment and climate as documented by long-term records of atmospheric constituents and climate (e.g. Keeling *et al.*, 1976; Thoning *et al.*, 1989; Hulme, 1995; Petit *et al.*, 1999; Brohan *et al.*, 2006). These changes provide a large-scale ecological experiment of the 'response' of land vegetation to external forcing. Such a response is not only of scientific interest per se but also of importance for predicting global carbon cycle feedbacks and in turn future greenhouse warming, and potentially changes in biodiversity. A system of particular importance in this regard is the forest vegetation of the Amazon basin as it accounts for a large fraction of the global land carbon store and biodiversity (e.g. Malhi *et al.*, 2007). Its vast size and largely nondeforested state makes this system also particularly suited to detect and study such responses.

In order to record such changes, a long-term network of permanent plots has been established in mature forests across Amazonia over recent years, uniting existing efforts of local botanists and foresters, known as 'RAINFOR' (Red Amazónica de Inventarios Forestales, or Amazon Forest-Inventory Network, http://rainfor.org, Malhi *et al.*, 2002), representing the com-

bined long-term ecological monitoring efforts of 35 institutions worldwide (Fig. 1). Measurements are mostly biometric, but by tracking the growth and death of individual trees, rates of change of aboveground biomass as well as mortality losses on a plot basis can be estimated.

For interpretation of such results the nature and representativeness of the statistical sample provided by the given network is paramount. On a basin-wide scale, plot locations have been selected to cover the main axes of variation of forest dynamics (soil fertility, precipitation, strength of El Niño signal). On a local scale forests have been selected which have not evidently been recently disturbed by human activities. However, large-scale natural disturbances thought to be mainly caused by intense wind gusts (Nelson *et al.*, 1994; Garstang *et al.*, 1998) cannot be avoided by simply choosing plots judiciously at a local scale. As a consequence the network provides a well-distributed sample of intact forest and associated small and large-scale disturbance regimes across the Amazon basin.

Analysis of earlier tropical plot data has suggested that large-scale changes in forest dynamics are currently occurring in Amazonia (Phillips & Gentry, 1994; Phillips *et al.*, 2004), and that an increase in aboveground biomass has occurred, with increases in mortality tend-
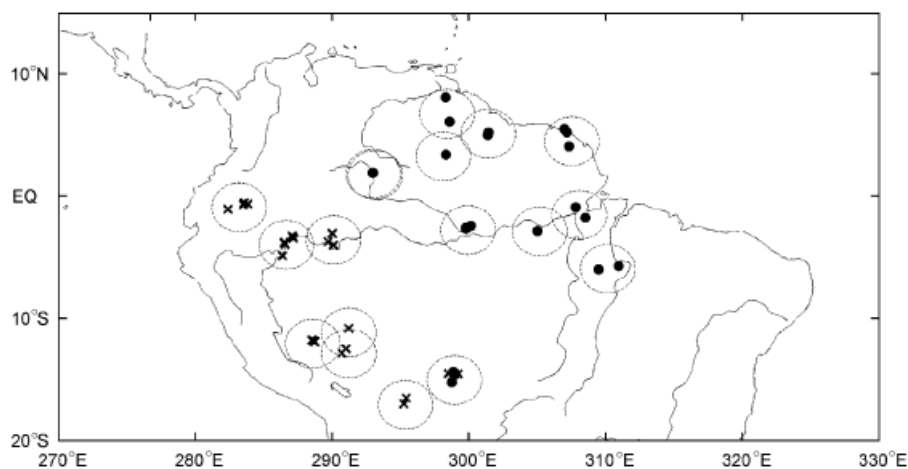

Fig. 1 Forest census plots from which data are used in this study. Crosses and dots indicate a rough categorization into Western and Eastern plots respectively (see text for details). The plots in Bolivia have been assigned to one or the other of the two groups according to geomorphology.

ing to lag increases in growth (Phillips *et al.*, 1998; Baker *et al.*, 2004a, b; Lewis *et al.*, 2004a). These conclusions have drawn major criticism of which the most significant is perhaps best characterized by the statement 'Slow in, Rapid out' (Körner, 2003). The 'Slow in, Rapid out' argument stresses that forest growth is a slow process while mortality can potentially be dramatic and singular in time, thereby entirely resetting forest stand structure almost instantaneously. As a consequence, sampling over comparably short observation periods may miss such more severe events. Inferences based on such sampling could therefore result in positively biased estimates of aboveground biomass trends in old-growth forests when results from a small network of large plots, or a large network of small plots are extrapolated to the whole basin (Fisher *et al.*, 2008).

In this paper, we address the hypothesis that observed biomass gains are indeed an artefact of insufficient spatio-temporal sampling. We proceed as follows: we first characterize growth and disturbance of the Amazon forests as recorded by the RAINFOR plots, using the term 'disturbance' here to describe any process associated with a decrease of living aboveground biomass (thus we do not distinguish between mortality due to senescence and external death processes). We then use the resulting disturbance frequency distribution to estimate the occurrence frequency of rare, large disturbance events over the last quarter century. As such large-scale disturbance events have not been recorded to date by the RAINFOR plots, it is unclear how to extrapolate the disturbance frequency distribution based on these data to large events. We therefore use two types of distributions for representing disturbances: a steeply decreasing distribution motivated by the RAINFOR statistics and more slowly decreasing distributions motivated by the Nelson *et al.* (1994) remotely sensed forest blow-down data. These distributions are then combined with observed distributions of growth into a simple stochastic simulator allowing us to study the statistics of aboveground biomass gains as a function of total observation period and plot ensemble size. We may then quantify the necessary sample size and time coverage to reduce biases due to the 'Slow in, Rapid out' character of forest dynamics. Finally we conclude with the implications for the robustness of the finding of increasing Amazon biomass in intact forests.

Materials and methods

Biometric field measurements

Results presented here are based on net changes in biomass (t ha^{-1} yr^{-1}) in forest inventory plots which in turn are the difference of two terms: biomass gains

(from tree growth and recruitment of new trees to the threshold size) and losses (from tree mortality) (Fig. 1). Measurement and analytical techniques have been described elsewhere (Baker *et al.*, 2004a, b; Lewis *et al.*, 2004b; Malhi *et al.*, 2004; Phillips *et al.*, 2004). Plots are typically 1 ha in size but frequently larger. We here analyse a total number of 135 plots with a total area of 226.2 ha for which the mean census interval is 3.2 years (standard deviation 2.8 years). Thus compared with the sample on which previous results of Baker *et al.* (2004a, b) were based, the sample size has approximately doubled. On average each plot has been censused 3.5 times, for a mean total observation period of 11.3 years. Aboveground biomass gains within a plot have been estimated based on measurements of tree diameter of all trees with diameter larger than 10 cm and biomass gains calculated from diameter increments using allometric equations derived from central Amazonia forests (Chambers *et al.*, 2001). These calculations include species-specific wood density values (Baker *et al.*, 2004b), and corrections for possible census-interval effects (Malhi *et al.*, 2004). Mortality rates of trees with diameter > 10 cm were determined by observation and where doubt existed by inspection of the cambium (wet or dry). Mortality rates have also been corrected for census–interval effects (Malhi *et al.*, 2004). We base our analysis on all censuses from the extended RAINFOR network starting as early as 1971 and concluding in 2006 after the 2005 drought.

Rare large-scale disturbance events and the power law

To assess robustness of conclusions drawn on modelling disturbance, we use a range of models for large-scale events bracketing existing observations. In particular, besides the RAINFOR data we base them on the only available dataset on large-scale disturbances. These disturbances attributed to high-intensity wind gusts have been compiled by Nelson *et al.* (1994) in Brazilian Amazonia, using remote sensing. In a recent paper Fisher *et al.* (2008) have proposed that the Nelson *et al.* (1994) blow-down frequency distribution for events occuring over one year follows a power law

$$p_{1\,year}(x) = \frac{\alpha - 1}{x_{\min}} \left(\frac{x}{x_{\min}} \right)^{-\alpha}, \qquad (1)$$

where α is the power law exponent and $x_{\min}$ the cutoff above which the power law is defined (e.g. Clauset *et al.*, 2007). They estimated the power law exponent using ordinary least squares (OLS) and found a value < 2. However, as demonstrated by Goldstein *et al.* (2004) the appropriate method for unbiased estimation of power law exponents is to use maximum likelihood estimation (MLE) (for completeness given in Appendix A1) as OLS

underestimates the exponent. Applying this methodology to the original Nelson *et al.* (1994) data, we find a power law exponent of 3.1, contrary to the results of Fisher *et al.* (2008).

To test in the following plausibility of theoretical distributions given data we use the bootstrapping method of Stute *et al.* (1993) which uses the Kolmogorov–Smirnov distribution distance metric $D = \max_{x_i}|F_{\text{data}}(x_i) - F_{\text{model}}(x_i)|$, $F_{\text{model}}(x) = \int_0^x p(x)\,dx$ where $F_{\text{data}}(x_i)$ is the empirical cumulated distribution function associated with the data $x_1, ..., x_n$ to estimate p values. p-values are defined as

$$p \equiv \frac{\#\ \text{bootstrap samples with}\ D_{\text{bootstrap}} > D_{\text{data}}}{\text{total}\ \#\ \text{bootstrap samples}}.$$

Thus, the closer p is to 1 the more plausible is the tested distribution given the data while $p < 0.1$ indicates that the theoretical distribution is not plausible given the data.

Relation between disturbance probability distributions from different census intervals

Data used in this analysis are based on a range of different census intervals. Thus, the question arises of how to combine these data, or in other words – what is the relation between empirical probability distributions based on different census intervals? We thus first demonstrate how these relations are established. If $p_{1\,\text{year}}(m)$ is the probability of a mass loss m due to mortality during a 1-year period then the probability for a biomass loss m during a 2-year period is the sum (integral) over all mass loss events m_1 during year 1 and m_2 during year 2 which add to a mass loss m over a 2-year period:

$$p_{2\,\text{year}}(m) = \int_{m_1 <= m} \underbrace{p_{1\,\text{year}}(m_1)}_{\substack{\text{mass loss}\\\text{during year 1}}} \underbrace{p_{1\,\text{year}}(m - m_1)}_{\substack{\text{mass loss}\\\text{during year 2}}} dm_1. \quad (2)$$

The assumption underlying this rationale is that mass loss events during subsequent years are independent from one another. If applied for example to an exponential distribution $p_{1\,\text{year}}(m) = \lambda e^{-\lambda m}$, then $p_{2\,\text{years}}(m) = (\lambda m)\lambda e^{-\lambda m}$. More generally for an n-year period

$$p_{n\,\text{years}}(m) = \int_{m_1 + m_2 + \cdots + m_{n-1} <= m} p_{1\,\text{year}}(m_1) p_{1\,\text{year}}(m_2) \cdots$$
$$p_{1\,\text{year}}(m - m_{n-1} - m_{n-2} - \ldots - m_1)$$
$$dm_1\,dm_2 \ldots dm_{n-1},$$
$$(3)$$

which for an exponential yields

$$p_{n\,\text{years}}(m) = \frac{(\lambda m)^{n-1}}{(n-1)!} \lambda e^{-\lambda m}. \quad (4)$$

The associated expectation value is n/λ in time units of $(n\ \text{year})^{-1}$ (and thus $1/\lambda$ when expressed in units of year^{-1}) and the variance is

$$\text{var} = \frac{1}{n\lambda^2}, \quad (5)$$

when expressed in units of year^{-2}. Similar results can be derived in principle from Eqn (2) for other distributions although for a power law they are somewhat complicated formulas including Euler Beta functions.

Disturbance severity and return time

The more severe a disturbance event the less often it is expected to be observed. But how rare are disturbance events of a given magnitude? We here propose that given an empirical probability distribution of biomass change (or mortality) per year, $p = p_{1\,\text{year}}(m)$, the relation between the severity of an event and its return time can be established as follows. The probability for a mass loss event with loss larger than m to occur per year is $P(X \geq m) = \int_m^\infty p(x)\,dx = 1 - \int_0^m p(x)\,dx = 1 - F(m)$ where $F(m) \equiv \int_0^m p(x)\,dx$ is the cumulative probability distribution function of the probability density $p(x)$. The return time τ of such an event therefore is

$$\tau \geq \frac{1}{P(X \geq m)} = \frac{1}{1 - F(m)}. \quad (6)$$

By inverting this relation, the biomass loss m associated with a given return time τ is then given by

$$m(\tau) = F^{-1}\left(1 - \frac{1}{\tau}\right). \quad (7)$$

As an example for an exponential probability density we obtain $\tau = \frac{1}{1 - (1e^{-\lambda m})} = e^{\lambda m}$ and $m(\tau) = \frac{\ln(\tau)}{\lambda}$, and for a power law $\tau = \left(\frac{m}{x_{\min}}\right)^{\alpha-1}$ and $m(\tau) = x_{\min}\tau^{1/(\alpha-1)}$, respectively.

Simple stochastic simulator of aboveground biomass balance

In order to establish the statistics of aboveground stock gains and losses implied by observed growth and mortality data, we formulated a stochastic simulator of the form

$$\frac{dM}{dt} = g - \mu, \quad (8)$$

where M is aboveground coarse woody biomass per unit area, t is time, g is a stochastic variable representing

the aboveground mass gain rate per unit area due to growth and μ a stochastic variable representing aboveground mass loss rate per unit area due to mortality which we subsequently simply term 'mortality'. To run the simulations, these parameters are estimated from the observed distributions as described in the following section, and we use a 1 year interval time step which is the natural choice given the observed disturbance statistics presented in 'Results and discussion'. To obtain the relevant statistics of dM/dt we repeated the stochastic simulator for an ensemble of 1000 virtual plots with each forest plot trajectory spanning a period of 100 years.

Because none of our plots has been affected by very rare large-scale blow-down events of the type observed by Nelson *et al.* (1994), we use in the following two variants of the observed disturbance parameterisation. Both variants follow observed exponential functions over the full range of aboveground biomass losses as observed by RAINFOR data, but differ in the way they treat larger and rarer events. Specifically, in one version the exponential distribution is assumed to extend 'ad infinitum' but for the second version, 'fat-tail' power law distributions replace the exponential decrease for the simulation of large disturbance events. The probability density for these mixed models is thus

$$p_{1\,year}(m) \propto \begin{cases} \lambda e^{-\lambda m}, & m < m_0, \\ m^{-\alpha}, & m \geq m_0. \end{cases} \quad (9)$$

Because the 'Slow in, Rapid Out' argument focuses on net system gains as opposed to balances of individual trees, the stochastic model is formulated on a plot level. Likewise, as this study is dedicated to existing data with the focus on aboveground biomass changes rather than individuals, we do not include any possible growth enhancement following disturbance. Nevertheless, the data do in fact show a small and relatively minor dependence of growth rates on disturbance, with growth rates slightly increasing after biomass loss. Inclusion of this effect, would tend to reduce the impact of individual disturbances on longer-term trends in stand biomass, meaning that our final conclusions on the robustness of growth rate trends inferred from RAINFOR plot data (Phillips *et al.*, 1998) are conservative.

Based on previous work indicating geographically determined differences among Amazonian forests in structure, dynamics, and floristics (Baker *et al.*, 2004a, b; Lewis *et al.*, 2004b; Malhi *et al.*, 2004; Phillips *et al.*, 2004; ter Steege *et al.*, 2006), standing stocks and biomass growth rate observations for plots across the Amazon basin were first clustered into two groups based on geographic location, i.e. 'West' (Peru, Ecuador,

Colombia, western Bolivia, Acre), and 'East' (Venezuela, Guyanas, Brazil except Acre). These macrogeographical categories correspond well to substrate age and soil fertility (Quesada *et al.*, 2008a, b). Forest plots in the geographically intermediate area of eastern Bolivia were allocated to 'West' and 'East' based on their nutrient status. Among Eastern Amazon plots standing stocks of biomass are markedly greater, and growth and mortality rates lower, than, Western Amazon forests (Fig. 2). Accordingly we parameterized stochastic processes separately for these two regions, but also undertook (combined) simulations for the Amazon as a whole.

Results and discussion

Mortality statistics

We characterized the mortality process by first developing histograms for census periods of suitable duration given the number of plot data available (Fig. 2). The
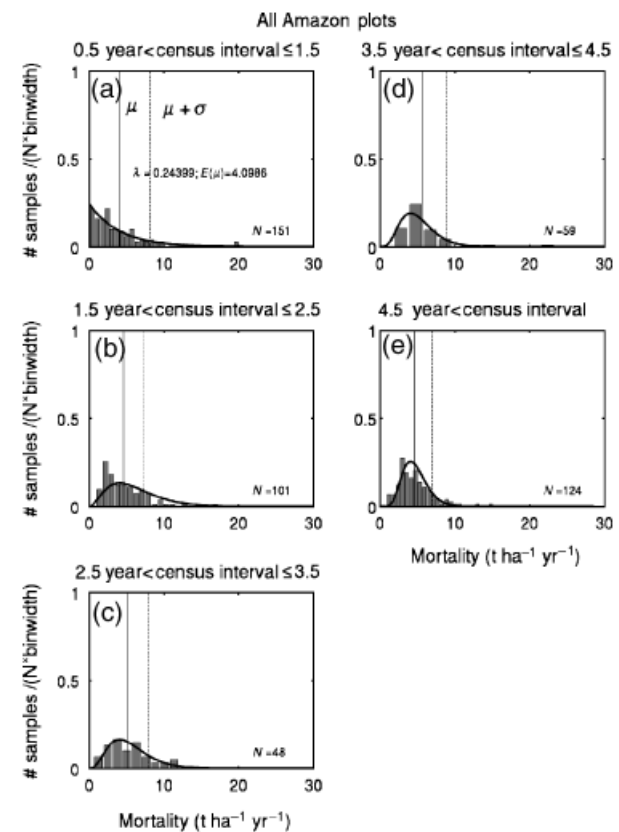

Fig. 2 Empirical distributions of decrease rate of living coarsewood aboveground biomass stocks due to mortality for increasing census interval lengths, maximum likelihood exponential distribution fit for (0.5, 1.5 years) interval and derived corresponding distributions for remaining intervals [cf. (Eqn 4)].

upper left panel of Fig. 2 shows the exponential distribution fit to the observed (0.5, 1.5 years) data obtained using the standard maximum likelihood estimator, the model curves for the other intervals being calculated from the fitted (0.5, 1.5 years) distribution using the formula for $p_{n\,years}(m)$ of Eqn (3). Agreement of the predicted distributions with the observed histograms is mostly very good [$p = 0.15, 0.79, 0.99, 0.82$ for $p_{1\,year}(m), \ldots, p_{4\,years}(m)$, respectively], confirming our simple rationale for inferring distributions from different census intervals. The comparably low p-value for the 1 year distribution is due to a poor fit for the very smallest disturbances. A better fit is obtained for a Weibull distribution ($p = 0.48$), which with exception of smallest disturbances, is nearly identical with the exponential distribution. The p-value for a power law is 0, indicating that this distribution is a poor descriptor of the mortality process. The good agreement between model predictions and data for multiyear periods also suggests that no essential mortality processes are being missed by using longer sampling intervals and that disturbance severity of subsequent years are nearly independent from one another.

The very good fit of the histogram for the (0.5, 1.5 years) interval over most of the observed range was confirmed by replotting the data histogram with axes scaled in various ways. For example, if the histogram of disturbance magnitudes does indeed obey an exponential distribution, then it should follow a straight line in a semi-logarithmic plot. On the other hand, if it follows a distribution with a fat tail (such as the power law function used by Fisher *et al.*, 2008), then it should follow a straight line in a full logarithmic plot. Figure 3 shows both types of plot, demonstrating that the data do follow an exponential function scaling relationship, but with some hint of a power law tail with exponent ~ 2 scaling the frequencies of the largest events. For the mixed exponential–power law distribution model we have therefore assumed power law tails with power law exponents of either $\alpha = 2$ or 3.1 (following the Nelson *et al.*, 1994 data) for $m > 25\,t\,ha^{-1}\,yr^{-1}$.

One predicted property of the modelled distributions is that the variance should decrease with increasing census interval as expected from Eqn (4) and this behaviour is indeed revealed by the data as well (Fig. 4). The fitted rate parameters λ for the exponential distribution as defined in Eqn (2) are 0.25, 0.22 and 0.25 $(t\,ha^{-1}\,yr^{-1})^{-1}$ for Eastern Amazon, Western Amazon, and all plots respectively. The similarity of parameters for the Eastern and Western Amazon gives, in hindsight, some justification for pooling plots to obtain a sufficiently large ensemble to characterize distur-
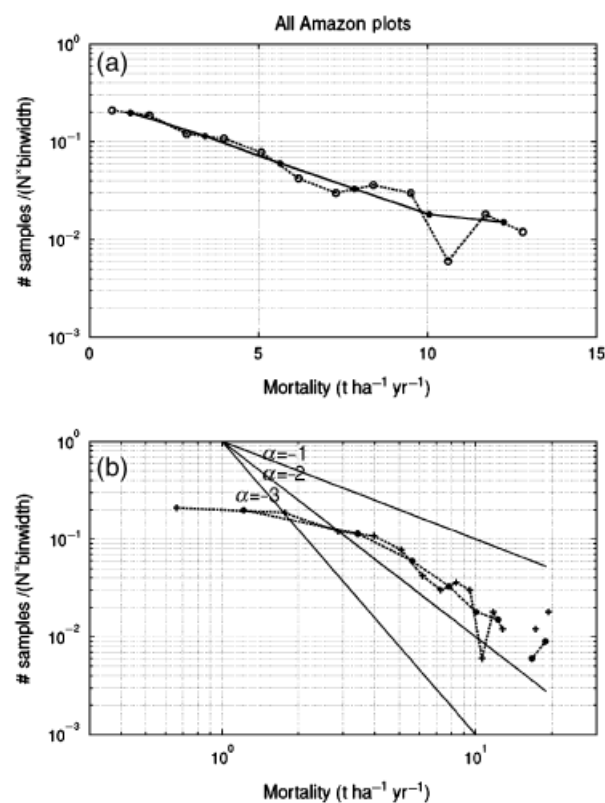

Fig. 3 Semi-logarithmic (upper panel) and logarithmic (lower panel) graph of empirical disturbance area distribution. The three lines in the lower panel indicate a range of power law distributions. Filled and open circles are based on different bin width used for calculating the histograms.

bance. It may potentially also tell us something about the mortality process itself. A speculative dimensional argument suggests

$$\lambda \sim \{\text{stem density}\,(ha^{-1})\}$$
$$\times \{\text{individual tree mortality rate}\,(yr^{-1})\}$$
$$\times \{\text{mean mass of individual}(t)\}$$

and therefore

$$(\text{stem density}) \times (\text{tree mortality rate})$$
$$\times (\text{mean mass of individual})$$
$$\sim \text{const}$$

across the basin. In order to assess the dependence of our results on plot size we have also repeated the same analysis but for plots with sizes between 0.5 and 1.5 ha only. The results were very similar (see Appendix A).

From the fitted distributions, we can infer how the likely frequency of severe disturbance events relates to their occurrence frequency using Eqn (6). Results are shown in Table 1 which shows, for example, that
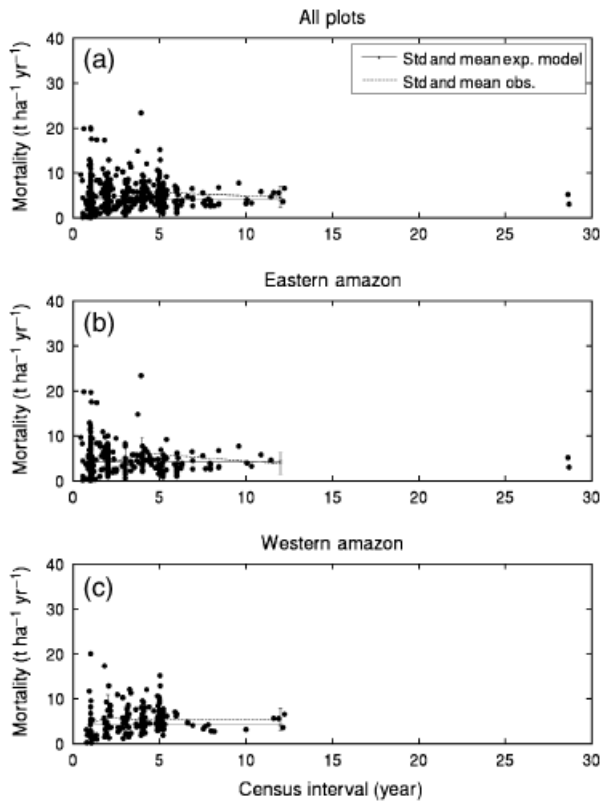
Fig. 4 Observed and predicted standard deviation [Eqn (4)] of empirical distributions as a function of census interval length.

Table 1 Predicted relation between severity of disturbance events and their return time

Return time τ (years)	Mortality loss $\geq$ (west/east) (t ha^{-1} yr^{-1})	Mortality loss $\geq$ (west/east) (%)
Exponential model		
20	12.5/13.6	5.0/3.6
100	19.1/20.9	7.6/5.5
200	22.0/24.1	8.8/6.3
1000	28.8/31.4	11.5/8.3
	All Amazon	All Amazon
*Mixed model**		
20	12.8	3.4
100	25.4	6.7
200	43.0	11.3
1000	96.4	25.4

*Exponential for $m < 25$ t ha^{-1} yr^{-1}, power law with exponent $\alpha = 2$ for $m \geq 25$ t ha^{-1} yr^{-1}.

according to the exponential model, events which remove 30 t ha^{-1} yr^{-1} should occur not more often than every 1000 years. However, according to the mixed model with power law exponent 2 which assumes a
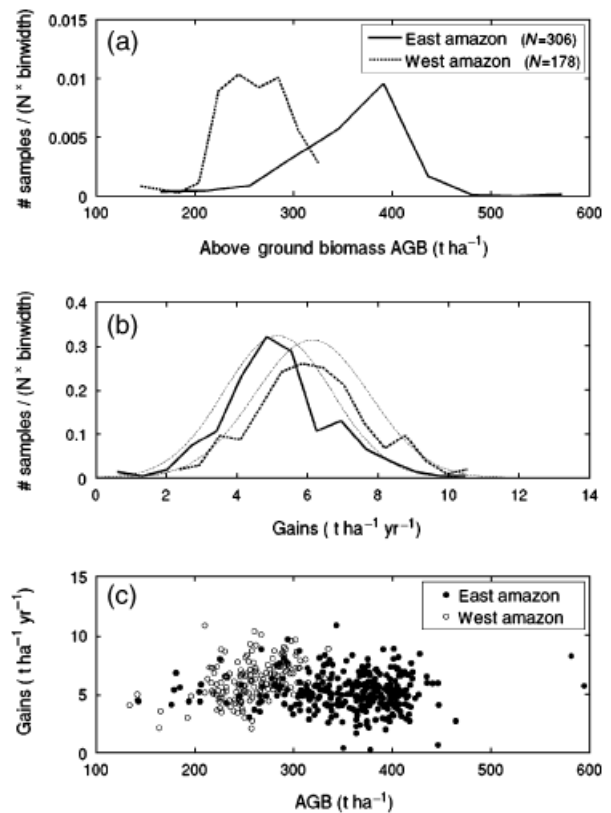

Fig. 5 Empirical distributions of living aboveground coarse-wood biomass stocks (a) (N is the number of censuses and bin width refers to the width of the bins used to calculate the histograms), aboveground coarse-wood biomass gains (b), and relation between the two (c) for Western and Eastern Amazon. Note that the gains displayed in (b) include only biomass increases. In contrast net biomass increase for a plot is given by the difference of gains minus losses.

more frequent high biomass disturbance regime they should occur approximately every 150 years. Nonetheless in both cases, we can still conclude that larger-scale disturbances as revealed by the RAINFOR network are very rare.

Modelling stand growth and biomass change

Since the data indicate there is nearly no functional dependence between gains and standing stocks (Fig. 5c), histograms of gains (Fig. 5b) indicate these can be approximated by a normal distribution $g \sim N(\mu, \sigma)$ with $\mu = 5.2$ (t ha^{-1} yr^{-1}), $\sigma = 1.5$ (t ha^{-1} yr^{-1}) for Eastern Amazonia and $\mu = 6.1$ (t ha^{-1} yr^{-1}), $\sigma = 1.6$ (t ha^{-1} yr^{-1}) for Western Amazonia respectively (parameters estimated using maximum likelihood and plausibility of distribution assessed by bootstrapping Stute *et al.*, 1993).
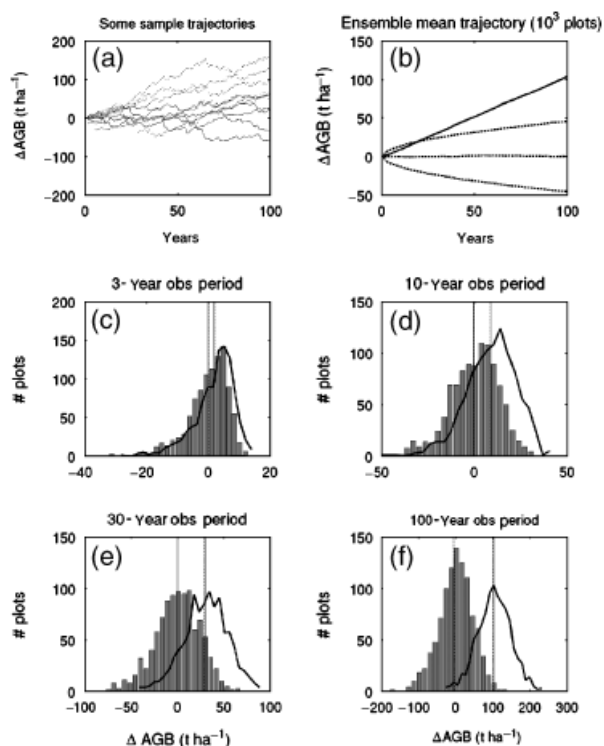
Fig. 6 Summary statistics of stochastic process predicting mass balance for a 1000 member sample for observed process characteristics (dots) and process adjusted such that gains balance observed mortality (line): a few members of the sample (a), time evolution of sample means and standard deviation (dotted) (b), and histograms of aboveground stocks changes ΔAGB for different observation from start of the process (c–f) with vertical lines indicating mean and median (dotted).

Combined with the mortality parameterization above, the growth function formulation g enabled a parameterization of the stochastic simulator [Eqn (8)] for two situations. Firstly, we ran the process for a system in equilibrium (light curve) by adjusting the centre point of the gains term such that it balances losses exactly when averaged over the 1000 plots over 100 years. These simulations provide us with the expected distribution from the 'null hypothesis' that the net biomass of Amazon forests is not actually increasing. Second, we utilized the observed stochastic characteristics of the gain term with sample trajectories shown in Fig. 6.

Figure 6c shows the statistical distribution of the integrated change in aboveground biomass (dM) expected for a 3-year observation period. Thus, even if the system is in equilibrium as a whole, the distribution of net biomass change is skewed as anticipated by Körner (2003) with the maximum (mode) of the distribution centered off zero. Another, second effect of the skewness of the 'null' distribution is an increase of the variance of the distribution compared with a normal distribution necessitating a somewhat larger sample of

plots to establish a statistically significant difference of the mean from zero. With an increasing period of monitoring the distribution loses its skewness, and tends towards a normal distribution, as expected.

When based on a 3-year observation period the equilibrium and nonequilibrium distributions are similar, but the distributions increasingly separate from each other as the observation period increases, with the data-based simulator indicating that after an observation period of a decade the net biomass change distribution has largely lost its skewness due to an insufficient sampling of the rare mortality events. Thus, to the extent that the data from the RAINFOR plots reflects the true statistics of Amazon forest disturbance as a whole, we suggest that a period of a decade is sufficient to circumvent this aspect of the 'Slow in, Rapid out' problem.

Number of plots required to detect signal

Using the distributions that we have found to be applicable to the modelling of Amazon forest dynamics, we can evaluate whether or not the number of plots in the network imparts sufficient statistical power for a verification of a net Amazon forest biomass gain occurring at the current time. Considering data from different plots and census intervals as independent estimates of net biomass gain rates the reasoning is as follows. The standard deviation of the statistical distribution underlying the net biomass gain rates sample is predicted by our stochastic simulator (Fig. 6b). The mean net biomass gain rate is then significant at the one sigma level if the standard deviation of the simulated mean is smaller than the observed mean, or, if the ratio between the standard deviation of the simulated mean and the observed mean is smaller than 1, and similarly for significance at the n-sigma level. The variance of the mean, as usual, scales inversely with the square root of the number of plots.

In order to establish significance according to this rationale it is helpful to notice that the variance of the distributions in Fig. 6b increases linearly with a slope of $20\ (t\,ha^{-1}\,yr^{-1})^2\,yr^{-1}$ for the exponential model, slope 22 $(t\,ha^{-1}\,yr^{-1})^2\,yr^{-1}$ for the mixed model with power law exponent 3.1 and slope 100 $(t\,ha^{-1}\,yr^{-1})^2\,yr^{-1}$ for the mixed model for power law exponent 2 (linear growth of the variance is equivalent to the standard deviation increasing with the square root of time which is what is shown in Fig. 6b as a dashed line). From this, the standard deviations of the distributions follow by calculating the square root (e.g. $\sigma(1\ year) = 4.6$ where σ is standard deviation of the distribution for the exponential model and $\sigma(1\ year) = 4.7$ and 10, respectively, for the mixed exponential–power-law models). As Table 2

Table 2 Summary of statistical significance of observed mean aboveground biomass gains

Census interval (years)	Eastern Amazon # censuses	r	Western Amazon # censuses	r	All Amazon # censuses	r
Exponential model						
(0.5–1.5)	121	0.40	27	0.87	148	0.37
(1.5–2.5)	80	0.35	21	0.67	101	0.31
(2.5–3.5)	18	0.60	30	0.47	48	0.37
(3.5–4.5)	24	0.45	35	0.38	59	0.39
(4.5–5.5)	21	0.44	49	0.28	70	0.24
			. . . . . .			
All*	303	0.19	178	0.19	481	0.14
*Mixed model*****						
All	303	0.41	178	0.41	481	0.30

Significance is assessed by the ratio $r \equiv (\sigma/\sqrt{N})/\Delta\overline{AGB}$ between standard deviation of the model based estimate of the sample mean $(\sigma/\sqrt{N})$ and the observed plot mean net biomass gain $\Delta\overline{AGB}$. N is number of censuses. Model predicted standard deviations σ for a 1-year period used for the table are 1.62, 1.55 and 1.63 t ha^{-1} yr^{-1} for the entire Amazon, Eastern Amazon and Western Amazon, respectively.

*Given that variance grows linearly with observation period and assuming independence of plot measurements we can scale variances to 1 year periods and use $\frac{1}{\sigma_{tot}^2} = \sum_i \frac{1}{\sigma_i^2}$ to estimate σ_{tot} for plots from different observation period lengths.

**Linearly proportional to exponential function for $m \leq 25$, and to power law function with $\alpha=2$ for $m \geq 25$ (t ha^{-1} yr^{-1}).

documents, given the data up to 2006 the claim of an increase in biomass over time is statistically robust at the 1σ level for all periods considered when using the exponential model and the mixed exponential power law model with power law exponent 3.1 (the standard deviations are nearly identical). However, it is not always robust at the 2σ level. The claim is significant at the 5σ level when all periods are combined both for the Eastern and Western Amazon.

If we repeat our analysis with the two mixed exponential–power-law models the main conclusions regarding significance of positive biomass gains remain robust, although the level of significance for the model with power law exponent 2 does decrease by approximately a factor two. It should be noted that the stochastic simulator based on the exponent 2 power law mixed model under-predicts net gains by a factor of three and thus is not supported by the observations (not shown). In contrast, the significance of the exponent 3.1 power law mixed model which is supported by the Nelson *et al.* (1994) data is actually virtually identical with the exponential model.

One may still argue that plots that are located close to one another do not provide spatially independent records. A rough and simple subjective assessment of the spatial distribution of plots (indicated by circles in the Fig. 1) indicates there are at least 17 spatially distinct clusters of plots Amazon-wide. This is a conservative assessment, because each grouping of plots is within itself greatly heterogeneous. Assuming temporal statistics to be decoupled from spatial statistics (P. Jansen *et al.*, in review) and using a mean census length of 3.2

years this reduces the effective census pool size to 39 for the entire Amazon. Using the same rationale as for creating the lowest line of Table 2 we obtain $(\sigma/\sqrt{N})/\Delta\overline{AGB} = 0.35$ for the entire Amazon. Thus, even when taking potential long-range correlations between plots into account then the conclusions regarding large-scale biomass gains across Amazonia remain significant at the two σ level.

Recently there has been a similar attempt to simulate the implications of such a sampling problem, but for a hypothetical sample based on only 1-year observation periods (Fisher *et al.*, 2008). This study likely overstates the 'Slow in, Rapid out' bias, when directly comparing the model and plot data results because the census interval length of the plot results is an order of magnitude greater than the census interval length of the modelling study. The Fisher *et al.* (2008) study also significantly underestimates power law exponents for a range of disturbance datasets (Lloyd *et al.*, unpublished data; Goldstein *et al.*, 2004; Clauset *et al.*, 2007). If the correct power law exponents had been used in their model, the results would similarly show that the results from the RAINFOR network were robust to this potential bias.

Summary

A network of long-term forest inventory plots across Amazonia shows, on average, a net increase in aboveground biomass. Given that additions of biomass from tree growth is approximately constant, yet, losses from mortality are occasional and stochastic, *a priori* we expect that our sample comes from a long-tail

distribution, if our sampling over time is shorter than the average time that forest plots take to recover from disturbance events. To explore this relationship we use a stochastic forest simulator, parameterized using the plot data, which shows that the distribution of net change in biomass is skewed for shorter intervals, as predicted. However, we show that the present-day sampling across Amazonia is sufficient to detect a positive trend in biomass over time. While there is little data with which to characterize the precise shape of the tail of the distribution, within the bounds of the available data, even if these occasional larger mortality events have, by chance, been under-sampled, they could not occur frequently enough to account for the increase in biomass seen across the network of inventory plots over the past 30 years.

Acknowledgements

References

Baker TR, Phillips OL, Malhi Y et al. (2004a) Increasing biomass in Amazonian forest plots. *Philosophical Transaction of Royal Society, Series B*, **359**, 353–365.

Baker TR, Phillips OL, Malhi Y et al. (2004b) Variation in wood density determines spatial patterns in Amazonian forest biomass. *Global Change Biology*, **10**, 545–562.

Brohan P, Kennedy JJ, Harris I et al. (2006) Uncertainty estimates in regional and global observed temperature changes: a new dataset from 1850. *Journal of Geophysical Research*, **111**, D12106, doi: 10.1029/2005JD006548.

Chambers JQ, dos Santos J, Ribeiro RJ et al. (2001) Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest. *Forest Ecology and Management*, **152**, 73–84.

Clauset A, Shalizi CT, Newman MEJ (2007) Power law distributions in empirical data. Available at http://arxiv.org/abs/0706.1062v2.

Fisher JI, Hurtt GC, Quinn Thomas R et al. (2008) Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecology Letters*, **11**, 554–563.

Garstang M, White S, Shugart HH et al. (1998) Convective cloud downdrafts as the cause of large blowdowns in the Amazon forest. *Meteorology and Atmospheric Physics*, **67**, 199–212.

Goldstein ML, Morris SA, Yen GG (2004) Problems with fitting to the power-law distribution. *European Physical Journal B*, **41**, 225–258.

Hulme M (1995) Estimating global changes in precipitation. *Weather*, **50**, 34–42.

Keeling CD, Bacastow RB, Bainbridge AE et al. (1976) Atmospheric carbon dioxide variations at Mauna Loa Observatory, Hawaii. *Tellus*, **28**, 538–551.

Körner C (2003) Slow in, Rapid out – Carbon Flux Studies and Kyoto Targets. *Science*, **300**, 1242–1243.

Lewis SL, Malhi Y, Phillips OL (2004a) Fingerprinting the impacts of global change on tropical forests. *Philosophical Transactions of the Royal Society, Series B*, **359**, 437–462.

Lewis SL, Phillips OL, Baker T et al. (2004b) Concerted changes in tropical forest structure and dynamics: evidence from 50 South American long-term plots. *Philosophical Transactions of the Royal Society, Series B*, **359**, 421–436.

Malhi Y, Baker TR, Phillips OL et al. (2004) The above-ground coarse woody productivity of 104 neotropical forest plots. *Global Change Biology*, **10**, 563–591.

Malhi Y, Phillips OL, Baker TR et al. (2002) An international network to understand the biomass and dynamics of Amazonian forests (RAINFOR). *Journal of Vegetation Science*, **13**, 439–450.

Malhi Y, Timmons Roberts J, Betts RA et al. (2007) Climate change, deforestation, and the fate of the Amazon. *Science*, **319**, 169–172, doi: 10.1126/science.1146961.

Nelson BW, Karpos V, Adams JB et al. (1994) Forest disturbance by large blowdowns in the Brazilian Amazon. *Ecology*, **75**, 853–858.

Petit JR, Jouzel J, Raynaud D et al. (1999) Climate and atmospheric history of the past 420 000 years from the Vostok ice core, Antarctica. *Nature*, **399**, 429–436.

Phillips OL, Baker TR, Arroyo L et al. (2004) Pattern and process in Amazon tree turnover, 1976–2001. *Philosophical Transaction of the Royal Society, B*, **359**, 381–407.

Phillips OL, Gentry AH (1994) Increasing turnover through time in tropical forests. *Science*, **263**, 954–958.

Phillips OL, Malhi Y, Higuchi N et al. (1998) Changes in the carbon balance of tropical forest: evidence from long-term plots. *Science*, **282**, 439–442.

Quesada CA, Lloyd J, Schwarz M et al. (2008a) Regional and large-scale patterns in Amazon forest structure and function are mediated by variations in soil physical and chemical properties. *Biogeosciences Discussions*, in press.

Quesada CA, Lloyd J, Schwarz M (2008b) Chemical and physical properties of Amazonian forest soils in relation to their genesis. *Biogeosciences Discussions*, in press.

Stute W, Gonzales Manteiga W, Quindimil MP (1993) Bootstrap based goodness-of-fit-tests. *Metrika*, **40**, 243–256.

ter Steege H, Pitman NCA, Phillips OL *et al.* (2006) Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, **443**, 444–447.

Thoning KW, Tans PP, Komhyr WD (1989) Atmospheric carbon dioxide at Mauna Loa Observatory 2. Analysis of the NOAA GMCC data, 1974–1985. *Journal of Geophysical Research*, **94**, 8549–8565.

Appendix A

Maximum Likelihood Estimation of (a) Power Law exponents and (b) parameters of multiyear convolutions of the exponential distribution

The likelihood L of model α given the data x is proportional to the probability of the data given the model. Thus for a power law,

$$L(\alpha|x) \propto p(x|\alpha) = \prod_{i=1}^{n} p(x_i|\alpha) = \prod_{i=1}^{n} \frac{\alpha - 1}{x_{\min}} \left(\frac{x_i}{x_{\min}} \right)^{-\alpha}$$

The most likely model is the one which maximizes L or equivalently its logarithm which, by setting its derivative equal to zero, yields,

$$\alpha = 1 + \frac{1}{\sum_{i=1}^{n} \ln\left(\frac{x_i}{x_{\min}} \right)}$$

(see e.g. Clauset *et al.*, 2007).

Application of the same principle to

$$p_{nyr}(m) = \frac{(\lambda m)^{n-1}}{(n-1)!} \lambda e^{-\lambda m}$$

yields as MLE $\lambda = n/\overline{x}$, where $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$. For evaluating goodness-of-fit of $p_{nyr}(m)$ to data using bootstrapping and MLE (Stute *et al.*, 1993), the cumulative distribution functions of $p_{nyr}(m)$ are needed. With integration by parts and by induction one finds

$$F_n(x) = \int_0^x p_{nyr}(x)dx = 1 - e^{-\lambda x} \sum_{k=0}^{n-2} \frac{(\lambda x)^k}{k!}.$$

Table A1 Summary of forest census plots used in this analysis. The data were extracted from the RAINFOR data base in July 2007

Plotname	Code	# plots	Latitude	Longitude	Large scale region	Total observation time (year)	Total area (ha)
Aguajal	AGJ	1	11 53 08.00 S	71 21 48.00 W	W	9.9	2.25
Altos de Maizal	ALM	1	11 48 00.00 S	71 28 00.00 W	W	10	2
Amacayacu: Agua Pudre	AGP	2	03 43 20.54 S	70 18 18.11 W	W	28.2	2
Amacayacu: Lorena	LOR	2	03 03 22.78 S	69 59 26.60 W	W	15.7	2
Allpahuayo	ALP	3	03 56 56.94 S	73 26 02.81 W	W	35.3	2.76
Añangu	ANN	1	00 32 00.00 S	76 26 00.00 W	W	4.9	1
Acuario	ACU	1	15 14 46.00 S	61 14 34.00 W	E	10.9	1
BDFFP	BDF	12	02 25 32.46 S	59 51 02.95 W	E	243.7	27
BEEM	BEE	2	16 32 00.00 S	64 35 00.00 W	W	8.5	2
Bionte	BNT	6	02 38 00.00 S	60 10 00.00 W	E	86.3	6
Bogi	BOG	2	00 41 54.66 S	76 28 55.86 W	W	23.2	2
Caxiuana	CAX	4	01 44 14.13 S	51 27 46.41 W	E	23.6	4
Chore	CHO	1	14 23 08.00 S	61 08 52.00 W	E	4.9	1
Cerro Pelao	CRP	2	14 32 16.90 S	61 30 01.22 W	W	14.4	2
Cuzco Amazonico	CUZ	4	12 29 56.34 S	68 58 25.63 W	W	68.9	4
El Dorado	ELD	1	06 06 07.63 N	61 24 12.12 W	E	32.2	1
Saint Elie	ELI	2	05 30 00.00 N	53 00 00.00 W	E	20.1	1.78
Forest Reserve Mabura Hills	FRM	1	05 13 12.00 N	58 34 48.00 W	E	6	1
Huanchaca Dos	HCC	2	14 33 39.00 S	60 44 55.00 W	W	20.3	2
Jacaranda	JAC	2	02 36 23.27 S	60 12 23.53 W	E	12	10
Jatun Sacha	JAS	6	01 04 06.00 S	77 36 55.00 W	W	72	5
Jenaro Herrera	JEN	2	04 52 41.12 S	73 37 46.02 W	W	4.1	2
Jari	JRI	1	00 53 40.00 S	52 11 25.00 W	E	11	1
Los Fierros	LFB	2	14 33 24.10 S	60 55 40.40 W	E	26	2
Las Londras	LSL	2	14 24 29.29 S	61 08 24.98 W	E	9.9	2

Continued

Table A1. (*Contd.*)

| Plotname	Code	# plots	Latitude	Longitude	Large scale region	Total observation time (year)	Total area (ha)
Manu	MNU	5	11 52 00.00 S	71 21 00.00 W	W	80.7	13.25
Maraba	MRB	3	05 40.00 S	49 02 00.00 W	E	23.3	6
Mishana	MSH	1	03 47 00.00 S	73 30 00.00 W	W	7.6	1
Nouragues	NOR	2	04 05 00.00 N	52 40 00.00 W	E	17.1	21
Pakitza	PAK	1	11 55 00.00 S	71 15 00.00 W	W	4	1
Paracou	PAR	10	05 15 00.00 N	52 50 00.00 W	E	100	42.4
Pibiri	PIB	3	05 01 18.39 N	58 37 15.96 W	E	39	3
Porongaba	RES	4	10 49 06.00 S	68 46 34.20 W	W	53.1	4
Roraima, Isla Maraca	ROM	1	03 25 00.00 N	61 40 00.00 W	E	11.5	2.25
Rio Grande	RIO	1	08 06 49.00 N	61 41 32.00 W	E	22.3	0.5
San Carlos de Rio Negro	SCR	4	01 55 58.22 N	67 01 18.06 W	E	75.6	4
Sacta	SCT	2	17 00 00.00 S	64 46 00.00 W	W	8.7	2
Sucusari	SUC	5	03 15 07.60 S	72 54 26.77 W	W	43.1	5
Tambopata	TAM	6	12 50 38.81 S	69 17 18.18 W	W	120.5	6
Tapajos	TAP	5	02 51 00.00 S	54 58 00.00 W	E	44	7.99
TEAM Caxiuana	TEC	5	01 42 23.51 S	51 27 33.28 W	E	18.5	6
TEAM Manaus	TEM	4	02 37 08.51 S	60 12 36.11 W	E	7.1	4
Tiputini	TIP	2	00 38 20.00 S	76 09 17.00 W	W	18.2	2
Yanamono	YAN	2	03 26 22.38 S	72 50 44.92 W	W	25.6	2
Zafire	ZAR	4	04 00 24.59 S	69 54 22.00 W	W	6.6	4
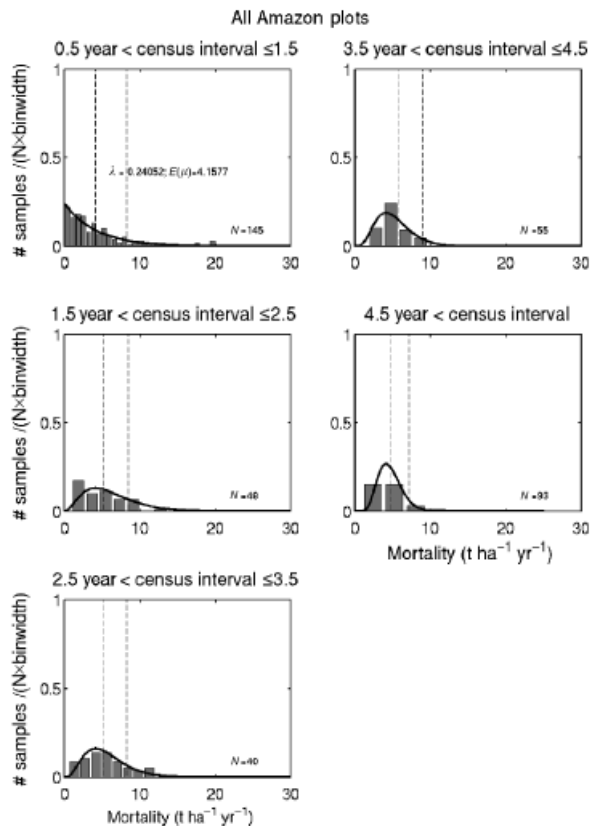
Fig A1 Dependence of disturbance statistics on size of plots. Same as Fig. 3 but with analysis restricted to 1 ha plots.

Table A2 Mortality histograms for all Amazon forest plots

ΔAGB (t ha^{-1} yr^{-1})	# Events
0.5 yr < Census Interval < 1.5 yr	
0.4427	22
1.1057	16
1.7686	16
2.4316	22
3.0946	10
3.7576	10
4.4206	9
5.0836	6
5.7465	10
6.4095	3
7.0725	4
7.7355	3
8.3985	4
9.0615	2
9.7245	3
10.3874	0
11.0504	2
11.7134	2
12.3764	1
13.0394	1
13.7024	0
14.3654	0
15.0283	0
15.6913	0
16.3543	0
17.0173	1

Continued

Table A2. (*Contd.*)

17.6803	1
18.3433	0
19.0063	0
19.6692	3

1.5 yr < Census Interval < 2.5 yr

1.3628	8
2.1782	21
2.9936	15
3.8091	10
4.6245	10
5.4399	9
6.2553	6
7.0708	8
7.8862	6
8.7016	1
9.5170	3
10.3325	1
11.1479	1
11.9633	0
12.7788	1
13.5942	0
14.4096	0
15.2250	0
16.0405	0
16.8559	1

2.5 yr < Census Interval < 3.5 yr

1.2185	4
2.4888	8
3.7592	10
5.0296	6
6.2999	9
7.5703	4
8.8407	2
10.1110	2
11.3814	3

3.5 yr < Census Interval < 4.5 yr

2.8296	12
4.7875	28
6.7455	11
8.7035	5
10.6614	0
12.6194	1
14.5773	1
16.5353	0
18.4932	0
20.4512	0
22.4092	1

4.5 yr < Census Interval

2.3396	50
5.1789	54
8.0182	14
10.8574	4
13.6967	2