Present and Future of Modeling Global Environmental Change: Toward Integrated Modeling, Eds., T. Matsuno and H. Kida, pp. 145–172.

Carbon Storage in the U.S. Caused by Land Use Change

Stephen W. PACALA[1], George C. HURTT[2], Paul R. MOORCROFT[1] and John P. CASPERSEN[1]

[1]*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, NJ 08540, U.S.A.*
[2]*Complex Systems Research Center, Institute for the Study of Earth, Oceans and Space, University of New Hampshire, Durham, NH 03824, U.S.A.*

Abstract—Here we examine the cause, size and future of the U.S. carbon sink. To estimate the size of the U.S. carbon sink we review a comprehensive land-based analysis of the carbon sink in the coterminous U.S. For the 1980s, the sink is between 1/3 and 2/3 PgC y^{-1}, and is split approximately evenly between forest and non-forest sectors. The nonforest sink is caused by fire suppression on non-forested lands, sediment burial in resevoirs, alluvium and colluvium, and agriculturual practices.

The forest sink has been attributed to changes in land use and the enhancement of plant growth by CO_2 fertilization, N deposition and climate change. To estimate the relative contribution of land use and growth enhancement in forest ecosystems, we use forest inventory data from five states spanning a latitudinal gradient in the eastern U.S. Land use is the dominant factor governing the rate of carbon accumulation in forests in these states, with growth enhancement contributing far less than previously reported. The estimated fraction of aboveground net ecosystem production due to growth enhancement is 2.0 +/– 4.4%, with the remainder due to land use.

To forecast the future of the U.S. carbon sink, we used the Ecosystem Demography Model (ED). We first modeled carbon sources and sinks from 1700–1990, and then projected patterns to 2100. Our projections indicate that the land-use portion of the U.S. carbon sink will decrease in the future, with a half-life of approximately 50 years, as U.S. ecosystems gradually equilibrate with current patterns of natural and anthropogenic disturbance.

INTRODUCTION

In this paper, we review and synthesize the results of four studies that we have recently completed with colleagues. All of the studies address the size and cause of the terrestrial carbon sink in the Northern Hemisphere.

For more than a decade, there has been widespread evidence for a terrestrial carbon sink averaging 1–2 PgC y^{-1} (10^{15} g carbon per year) in the Northern Hemisphere (1). However, the cause of the sink remains controversial. In particular, one study estimated that the majority of the sink was in North America during 1988–1992 (2), while other studies estimate no North American bias (3).

Inventories of terrestrial carbon storage in the coterminous United States (the U.S. minus Alaska and Hawaii) appear to support the conclusion that the sink is small (4). However, inventories in the Northern Hemisphere have been able to account for only a third or less of the 1–2 PgC y^{-1} indicated strongly by other lines of evidence (5). Moreover, the two primary groups of U.S. inventory studies strongly disagree about cause of the U.S. sink. Houghton and his colleagues (6) used historical records of land use change, timber production, soil conservation, wildfire rates, and simple models of carbon gains and losses in vegetation, soils and wood products. They estimated a sink for the coterminous U.S. averaging 0.39 PgC y^{-1} from 1950–1990, and caused primarily by increases in crop productivity and changes in the management of agricultural soils (0.15 PgC), and by fire suppression on non-forested land (0.13 Pg C). They concluded that the entire forest sector contributed only 0.07–0.12 PgC y^{-1}.

The U.S. Forest Service (USFS) estimated a coterminous U.S. sink averaging 0.33 PgC y^{-1} from 1952–1992, using census and tree measurement data from the over 100,000 plots in their Forest Inventory and Analysis (FIA) network, together with models of soil carbon and the fate of wood products (7). Although 0.33 Pg is close to 0.39 Pg, the entire USFS estimate is for the forest sector. If all of the carbon identified in both (6) and (7) were real, then the annual sink for the coterminous U.S. would be 0.60–0.65 PgC (0.33 from (7) plus 0.39 from (6) minus the forest sector estimates from (6)) and thus the overlap between the two estimates is only 11–20% of the total (0.07/0.65 to 0.12/0.60).

The FIA data base shows that the increase of carbon in trees has remained remarkably steady from 1952 to 1992, at approximately 0.10 + 0.02 PgC y^{-1}, because regrowth in the eastern half of the U.S. consistently exceeds harvest by about 0.1 PgC (7). This value is approximately double the increase in living forest carbon modeled in (6). To first order, differences among published inventories based on FIA data are caused by differences in the modeling of all forms of dead organic matter (including slash, wood products, standing dead trees, and soil carbon). For example, one study produced an estimate of only 0.08 PgC y^{-1} for the 1980's, because its modeling assumptions led to negligible accumulation of nonliving carbon (8). In contrast, the assumptions behind the USFS estimate of 0.33 PgC y^{-1} implied that soil carbon accumulated twice as fast as living carbon in trees. In addition, no comprehensive inventory has as yet included the carbon sink caused by sediment burial in reservoirs, alluvium and colluvium and by the transport of carbon into the oceans by rivers. At least one study suggests that sediment burial and river transport may be significant (9). Finally, no comprehensive inventory accounts for net export of carbon in agricultural and wood products.

Ecosystem models provide the final source of information about the terrestrial carbon sink and generally produce small estimates for the coterminous U.S. For example, the models in the recently published VEMAP comparison produced estimates of 0.08 + 0.02 PgC y^{-1} for the period from 1980–1993 (10). However, none of these models includes the land use changes (i.e. agricultural abandonment, fire suppression, forest harvesting and regrowth, no-till agriculture) that play

such a dominant role in the inventory analyses. The models focus instead on the effects of climate change and CO_2 and nitrogen fertilization.

In what follows, we first summarize the results of a new inventory-based analysis of the coterminous U.S. carbon sink, which shows that the sink averages between one third and two thirds Pg annually (11). These estimates and smaller than the 0.81–0.84 PgC y^{-1} in the controversial study (2) (see 11 for a discussion of the portion of estimates in (2) that correspond to the coterminous U.S.). However, they are significantly larger than the previously published range (one tenth to one third PgC y^{-1}). The analysis in (11) shows that approximately half the sink can be unambiguously ascribed to land use and management. We then summarize a recently published analysis (12) of the FIA data showing that the other half of the U.S. carbon sink is also overwhelmingly caused by land use and management.

Given that human land use rather than CO_2 or nitrogen fertilization or climate change causes the sink, it is interesting to consider the sink's future. We then turn to two additional studies. The first (13), introduces a new model that incorporates the sub grid-scale heterogeneity necessary to simulate land use. The second (14), applies this model for the past 300 years of land use in the coterminous U.S., and shows that the U.S. sink will decrease throughout the coming century. Unlike sinks caused by fertilization or climate change that might increase, the land use sink will decrease as U.S. ecosystems adjust to the altered disturbance regimes created by land use and management.

STUDY 1—NEW LAND-BASED ESTIMATES OF THE CARBON SINK IN THE COTERMINOUS U.S. (PACALA *ET AL*., 2001)

Table 1 contains upper and lower bounds for seven factors contributing to the carbon sink in the coterminous U.S. carbon sink during the 1980's. The lower bound for forest trees is the USFS estimate (7), while the upper bound reflects uncertainty in the allometric equations that convert the raw diameter data into biomasses (11).

The FIA census program has not historically included systematic measurements of litter, woody debris, slash and mineral soil, and recent expansions of the program have not been in place long enough to obtain direct estimates of carbon fluxes. For this reason, land-based analyses rely on models to estimate changes in non-living forest carbon, given the measured age structure and productivity of forest stands, and the history of land use and fire suppression. The upper bound for "Other Organic Matter in Forest" in the table comes from the study (14) discussed below. The lower bound combines convergent estimates for the accumulation of slash and woody debris (6, 7), with minimum estimates for the accumulation of soil carbon (15, 16). Note that published estimates lower than the lower bound in the table or higher than the upper bound have now been revised by their primary authors in the study (11).

Wood products create a carbon sink because they accumulate both in use and in landfills. The bounds in Table 1 are the lowest and highest numbers published (7, 17).

Table 1. Sinks of carbon for 1980–1990 in the coterminous United States (PgC y^{-1}).

Category	Low	High
Forest trees	0.11	0.15
Other organic matter in forests	0.03	0.15
Domestic wood products	0.03	0.07
Woody encroachment on non-forested lands*	0.12*	0.13*
Agricultural soils	0.00	0.04
Exports minus imports of food and wood products	0.04	0.09
Sediment burial, and river export	0.04	0.08
Apparent** U.S. Sink	0.37	0.71
Sink***	0.30	0.58

*These numbers are not bounds, but rather the only two existing estimates.

**By "apparent" sink we mean the net flux from the atmosphere to the land that would be estimated in an inversion. It includes all terms in the table.

***Excludes sinks caused by the export/import imbalance for food and wood products and river exports (0.03–0.04 PgC y^{-1} of the Sediment Burial and River Export term) because these create corresponding sources outside the U.S.

Over a large part of the western region, woody plants were historically excluded by recurrent fire and are now increasing, with corresponding increases in carbon storage. Because the extent of woody encroachment is not known, we cannot estimate reliable upper and lower bounds. The values in Table 1 represent the only two large-scale estimates that we know of for woody encroachment on nonforested lands in the coterminous U.S. (6, 14). It is important to understand that these numbers are not bounds in the same sense as the other numbers in the table. Of the seven separate items in the budget, there is a substantial chance that the correct value lies outside the reported bounds only in the case of woody encroachment.

Recent management practices that may have led to increases in carbon storage on U.S. agricultural soils include the Conservation Reserve Program (i.e. conversion of unproductive croplands to perennial grassland), expanded use of no-till agriculture (in which seeds are drilled into the ground), and improved productivity caused by new plant varieties and increased fertilizer inputs (6). Field and modeling studies suggest that U.S. agriculture represented a small carbon source during 20th century until approximately 1970–1980, and then became a small sink. The bound in the table are from recent modeling studies with the CENTURY model (18, 10).

The U.S. exports more carbon than it imports. This imbalance is dominated by agricultural products with only a minor contribution from wood products. The bounds in the table were calculated from FAO (19) and USDA (20) statistics (11). It is important to understand that this term represents a net flux from the atmosphere to the land, but not a global sink for atmospheric carbon, because most of the agricultural imbalance returns to the atmosphere elsewhere.

The final term in the table is derived from extensive data on reservoirs and rivers, extrapolated by models of sediment transport, deposition, erosion and mineralization (21, 11). One quarter to one half of this flux is caused by burial of eroded sediments containing carbon in U.S. reservoirs, alluvium and colluvium. The remainder is caused by export of carbon by rivers into the sea. Note that most of the carbon exported into the sea will return to the atmosphere by oceanic processes, and thus contributes only the local U.S. sink, and not to the global sink.

Collectively the bounds in Table 1 show that the conterminous U.S. carbon sink averaged roughly 0.5 + 0.2 PgC y^{-1} during the 1980s, with only about one half of the sink caused by accumulating carbon in the forest sector.

STUDY 2—THE SINK IS CAUSED BY LAND USE CHANGE (CASPERSEN *ET AL*., 2000)

While there is a convergence among estimates of the size of the U.S. carbon sink, considerable uncertainty remains about the cause of the sink. The inventory results suggest that the carbon accumulation in terrestrial ecosystems may be explained in large part by changes in land-use, such as fire suppression and agricultural land abandonment. However, carbon accumulation may also be explained by the enhancement of plant growth CO_2 fertilization, nitrogen deposition and climate change. The challenge is to evaluate the relative contribution of land-use and growth enhancement.

Here, we examine whether growth enhancement has increased the rate of biomass accumulation in eastern U.S. forests using the Forest Inventory and Analysis (FIA) database (22–24). We restrict our analysis to five eastern states (Minnesota, Michigan, Virginia, North Carolina and Florida) which provide the data necessary to estimate changes in biomass (12). For each of the more than 20,000 plots in the five-state sample, the data include estimates of aboveground biomass at the time of two successive inventories (the first in the late 1970's to mid 1980's and the second in the early to mid 1990's), changes in aboveground biomass due to growth and mortality (25), and the age of the plot in each of the two inventory periods (defined as the time since stand establishment following agricultural abandonment, clearcutting, or stand-destroying natural disturbance).

The best way to estimate historical changes in growth and mortality rates (hereafter referred to as vital rates) is with repeated measurements spanning the period of interest. However, this would require at least three inventories. How then can we estimate historical changes in vital rates with only a single measurement provided by the two inventories? The key is that in addition to stand age we also have a record of stand biomass which is the cumulative result of past vital rates. The biomass B of a stand of age A is the sum of its gains from growth minus its losses from mortality over the previous A years. Therefore, if vital rates measured in the 1980's were unchanged from the past, we could predict the biomass of a stand of age A by calculating the difference between growth and mortality at each previous age, using the 1980's rates, and summing the difference over the previous A years. However, if growth rates were higher in the 1980's than in the

past, or mortality rates were lower, then the biomass predicted using 1980's rates would be larger than observed. Thus, we can use the difference between the predicted and observed biomass to estimate past changes in the vital rates.

To illustrate this method, we present a deliberately simplified example using the Michigan inventory data. The first step is to estimate the current growth, $G(A)$, and the current mortality rate, μ ($\mu B = M$, where M is the total amount of biomass lost to mortality) (26). Figure 1 shows that growth in Michigan in the 1980's increases rapidly with stand age as the canopy closes, and then remains approximately constant at about 2.4 tons $ha^{-1}yr^{-1}$. The mortality rate in Michigan (the fraction of biomass lost to mortality, including both natural mortality and selective harvesting) is 0.021 yr^{-1}. Assuming that past vital rates were the same as the vital rates in the 1980's, we can calculate the expected relationship between biomass and stand age, $B(A)$, using $G(A)$ and μ. For example, if $B(0)$ is the initial biomass (the biomass present following clearcutting, natural stand destruction, or agricultural abandonment), then the biomass at any subsequent age can be calculated interatively:

$$B(1) = B(0) + G(0) - \mu B(0)$$

$$B(2) = B(1) + G(1) - \mu B(1)$$

etc.

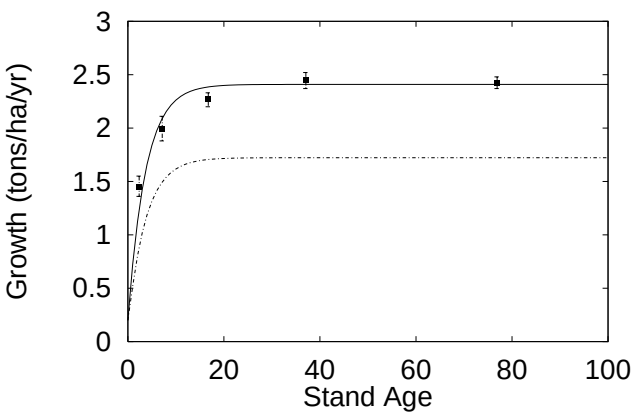

Fig. 1. The 1980's growth rate in Michigan (squares). Mean and 1 SE of the data are shown for five different age class bins. The small standard errors reflect the large sample size (2,890 plots). The upper curve (solid line) is referred to as *G(A)*. If past growth rates were lower than the 1980's growth rate, then forest biomass would be lower than is observed. To illustrate we present a hypothetic scenario in which we assume that growth prior to 1930 was given by the lower curve (dashed) and then increased linearly ($\beta = 0.005$) until it reached the upper curve in Fig. 1 in 1980. The amount of biomass predicted to accumulate under the hypothetical scenario appears in Fig. 2.

The biomass predicted from 1980's rates ($B(A)_{pred}$) is given by the upper curve in Fig. 2. Note that the vital rates measured in the 1980's accurately predict the observed biomass in the 1980's ($B(A)_{pred} \approx B(A)_{obs}$), even though the observed biomass is the cumulative result of vital rates prior to the 1980's. The implication is that the vital rates in the 1980's were not substantially different from historical vital rates, and therefore there has been little or no growth enhancement.

To estimate the magnitude of growth enhancement, we use the same iterative procedure for summing past vital rates, except we allow past growth rates to vary with time to obtain the best fit to the observed relationship between biomass and stand age:

$$B(1) = B(0) + G(0)(1 - \beta) - \mu B(0)$$

$$B(2) = B(1) + G(1)(1 - \beta) - \mu B(1)$$

etc.

Here, $G(A)$ is the growth rate of an age-A stand and $G(A)f(t)$ is the growth rate of an age-A stand t years ago. We estimate β as the value that provides the best

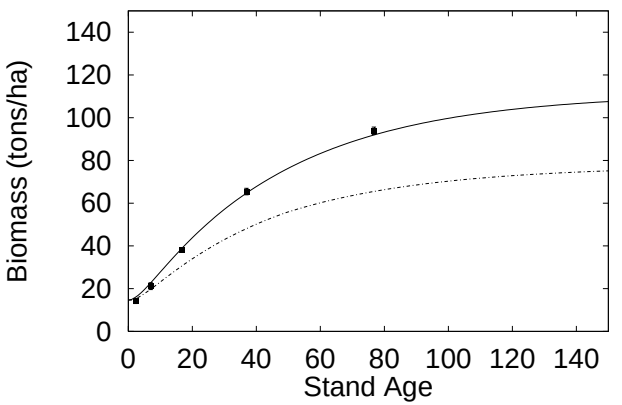

Fig. 2. Observed biomass ($B(A)_{obs}$) in Michigan in 1980 (squares). Mean and 1 SE are shown. $B(A)_{pred}$ (solid line), the amount of biomass predicted to accumulate at current growth and mortality rates, assumes no growth enhancement. Note that the current mortality rate includes both natural mortality and selective harvesting, so the asymptotic biomass is lower than would be expected in undisturbed old-growth forests. The amount of biomass predicted to accumulate under the hypothetical scenario (dashed line) assumes that growth prior to 1930 was given by the lower curve in Fig. 1 and then increased linearly ($\beta = 0.005$) until it reached the upper curve in Fig. 1 in 1980. The hypothetical accumulated biomass lies many standard errors below the observed biomass because the hypothesis that growth was lower in the past is not consistent with the observed biomass. We chose this scenario because, if it were true, then 50% of total forest biomass accumulation in Michigan during the 1980's would have been due to growth enhancement. The implication is that the method provides the sensitivity necessary to assess the magnitude of growth enhancement.

fit to $B(A)_{obs}$. The estimated value of β is very small ($\beta < 0.00001$, indicating a growth enhancement of less than 0.001% per year) because $B(A)_{pred} \approx B(A)_{obs}$. In contrast, if past growth rates were significantly lower than the current growth rate, then forest biomass would be significantly lower than observed (see the hypothetical scenario presented in Figs. 1 and 2 in which growth enhancement is assumed to be large enough to cause 50% of current carbon accumulation).

While the analysis presented above illustrates how in principle the magnitude of growth enhancement can be estimated from inventory data, there are two issues which must be addressed to make the analysis rigorous. First, we must account for the strong correlations that develop between growth, mortality and biomass due to spatial and temporal variation in growth and mortality. Second, we must allow for the possibility that a difference between predicted and observed biomass could be due to changes in mortality as well as changes in growth. To address these issues we derived a maximum likelihood estimator that allows for correlated errors among $B(A)$, G and M and changes in mortality as well as growth (11).

To estimate the contribution of growth enhancement to carbon accumulation, we estimated a β for each state and calculated the fraction of above-ground net ecosystem production (ANEP) due to growth enhancement. First we calculated ANEP ($B(A + 1) - B(A)$ with the estimated growth enhancement ($ANEP_{with}$) and also with growth remaining constant after 1900 ($ANEP_{without}$). Then, the fraction due to growth enhancement was calculated as $1.0 - ANEP_{without}/ANEP_{with}$. The fraction of ANEP due to growth enhancement since 1900 is 2.0 +/– 4.4%, with the remainder due to land-use history.

Because the results demonstrate that carbon accumulation is overwhelmingly due to forest regrowth rather than growth enhancement, we conclude that land-

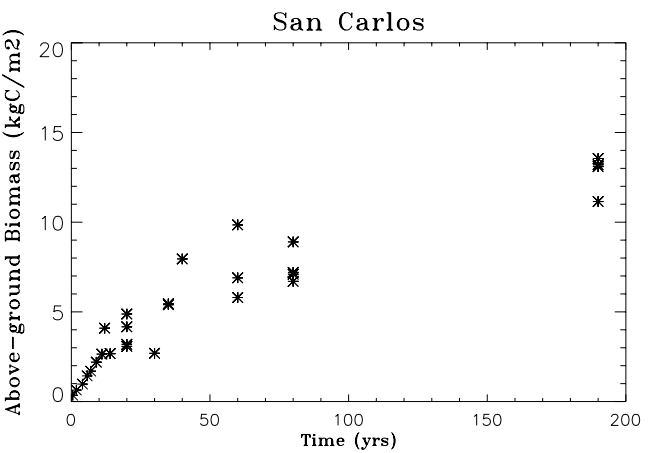

Fig. 3. Chronosequence of above-ground biomass (kg-C m^{-2}) at San Carlos (2°N, 68°W), a tropical forest site recorded by Uhl (28) and Saldarriaga *et al.* (27).

use change is the dominant factor governing the rate of carbon accumulation in eastern U.S. forests. It is possible that growth enhancement by N deposition and CO_2 fertilization is balanced by the negative effects of other factors such as ozone and calcium depletion. However, such balancing effects on growth cannot be readily evaluated with the FIA data. Our results have several implications for the effort to predict and model the fate of the terrestrial carbon sink. First, since regrowth accounts for the vast bulk of the carbon accumulating in the forests we studied, carbon accumulation in forests may be expected to attenuate at a predictable rate as forests recover their former biomass. Second, ecosystem models which focus exclusively on physiological processes omit the dominant factor governing the rate of carbon sequestration in forests, namely land-use history. Finally, ecosystem models which account for land-use history, stand age, and stand structure may be expected to predict gross carbon fluxes in forests despite considerable uncertainty regarding the effects of N deposition, CO_2 fertilization and climate change.

STUDIES 3 AND 4—MODELING THE FATE OF LAND USE SINKS (MOORCROFT *ET AL*., 2001 AND HURTT *ET AL*., IN REVIEW)

Ecosystem Demography model (ED)

The problem of scale has been a critical impediment to incorporating important fine-scale physiological and ecological processes into global ecosystem models. We have addressed this issue by developing a new, terrestrial biosphere model, the Ecosystem Demography model (ED), that incorporating a general method for scaling stochastic, individual-based models of vegetation dynamics to large-scales. ED is formulated as an individual-based, stochastic plant simulator but uses a size- and age-structured (SAS) approximation to predict its large-scale ensemble mean behavior.

Figure 3 shows the accumulation of above-ground carbon in a moist tropical forest in Venezuela (27) as a function of successional age (time since abandonment of slash-burn agriculture). The greater than 200 year timescale of uptake is surprisingly long, especially because net primary productivity (NPP) is high throughout the chronosequence shown, 1.1–1.2 kg-C $m^{-2}yr^{-1}$ (28), thus ruling out the time required to rebuild depleted soil fertility as an explanation for the length of the timescale. With more than half of a tree's NPP going to wood production (29, 27), one would expect that the time required to accumulate 12 kg-C m^{-2} above ground would be an order of magnitude shorter than is observed. Indeed, most global ecosystem models would predict a short timescale for this process (30–32).

The detailed surveys of the plots at San Carlos provide an explanation for the long timescale of carbon uptake (27, 28). As we show below, the rate of carbon accumulation at San Carlos is explained by height-structured competition, the presence of successional diversity, and the demography of size and age distributions. Because of the difficulty of phenomenologically parameterizing the outcome of size- and age-structured competition between plants of different

types under different environmental conditions, size- and age-structured ecosystem models may be necessary to predict transient carbon dynamics and fluxes.

Individual-based models of vegetation dynamics (gap models)

A gap model is a stochastic process that predicts the fate of every individual inhabiting an area the size of a canopy tree. This area is labeled a "gap" and within each gap individuals in the gap compete for light, water and nutrients. The model is stochastic because of the stochastic nature of birth, death and dispersal. The success of gap models derives from the small scale of their formulation (33). These models naturally capture the gap-scale heterogeneity created by the deaths of single canopy trees, and the height-structured competition among saplings competing to fill an opening in the canopy. Because the models are formulated at the scale at which field biologists work, it is comparatively straightforward to measure relevant parameters (34) and compare to data on forest structure (33).

The development of global gap models has largely been prevented by the difficulty of predicting their large-scale behavior. Predictions at scales of forest stands or larger are made by simulating the stochastic process for an ensemble of coupled or uncoupled gaps which are then summed or, equivalently, averaged. While this brute force simulation approach has been successfully used at hectare and kilometer scales it is not been feasible at larger scales (for example a 1°GCM grid cell would contain between 10^8 and 10^9 modeled gaps). What is needed to scale from gap to global dynamics is a way to derive the equations governing the ensemble average of a stochastic gap model directly from the fine-scale processes in it. In this section, we outline the components of the ED model, then derive a set of partial differential equations that govern its ensemble average behavior. These equations scale the processes in the gap model and represent a form of ecological "statistical mechanics".

Overview of the ED model

ED is formulated as a physiologically-driven, individual-based, stochastic plant simulator driven by climate and soil inputs with associated models of below-ground carbon, water and nitrogen dynamics (Fig. 4). A detailed description of the ED model can be found in (35). The model is implemented as a stochastic process in a series of simple steps. Suppose that a simulation of Q gaps (y = 1, 2, ..., Q) that share a common set of climate and soil inputs is currently at time t and that we wish to produce the state of the system at time $t + \Delta t$. First, we grow each individual's structural stem in size by an amount $g_s(\underline{z}, \underline{x}, \underline{r}, t)\Delta t$ and each plant's living biomass by an amount $g_a(\underline{z}, \underline{x}, \underline{r}, t)\Delta t$.

Here $\underline{r}$ is shorthand for a vector $\underline{r}(h, y, t)$ of local resource levels, which has three elements: light, water and nitrogen. The growth functions integrate the output of the leaf-level physiological model given local light, water and nitrogen availability, transforming them into growth of structural and active tissues using an allometric sub-models. The plant's size is given by a two dimensional vector

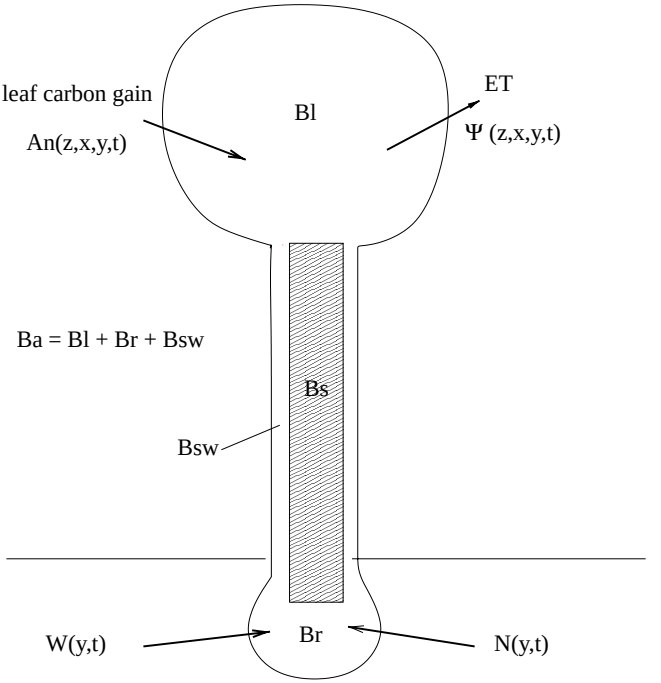

Fig. 4. Individual-level fluxes of carbon, water and nitrogen and the partitioning of carbon between active and structural tissues (B_a and B_s respectively).

$\underline{z} = [B_s, B_a]$ because each plant is composed of structural stem tissue B_s and active (leaves, sapwood and fine root tissue) B_a (Fig. 4), and the height and diameter of the individual is uniquely related to B_s. Second, the mortality sub-model determines each plant's probability of mortality, $\mu(\underline{z}, \underline{x}, \underline{r}, t)\Delta t$, killing the plant if a simulated pseudo-random coin toss with this probability indicates death. Third, each plant gives birth to an offspring with probability $f(\underline{z}, \underline{x}, \underline{r}, t)\Delta t$ and new recruits are assigned to gaps at random. Fourth, we use a fire sub-model to calculate the probability of fire, $\lambda_F(y, t)\Delta t$, for each gap and toss simulated pseudo-random coins to determine which gaps burn. Finally, we calculate, for each gap, the changes in the below-ground amount of water, soil carbon and soil nitrogen, using hydrological and decomposition sub-models.

Figure 5(a) shows an example of output from the stochastic simulator for 25 linked gaps at San Carlos, a wet rainforest site. The output is obtained by using the above algorithm to simulate the carbon and nitrogen capture, water loss, growth, reproduction, and mortality of every individual throughout its life cycle and the associated below-ground dynamics of carbon, nitrogen and water within each gap in the simulator.

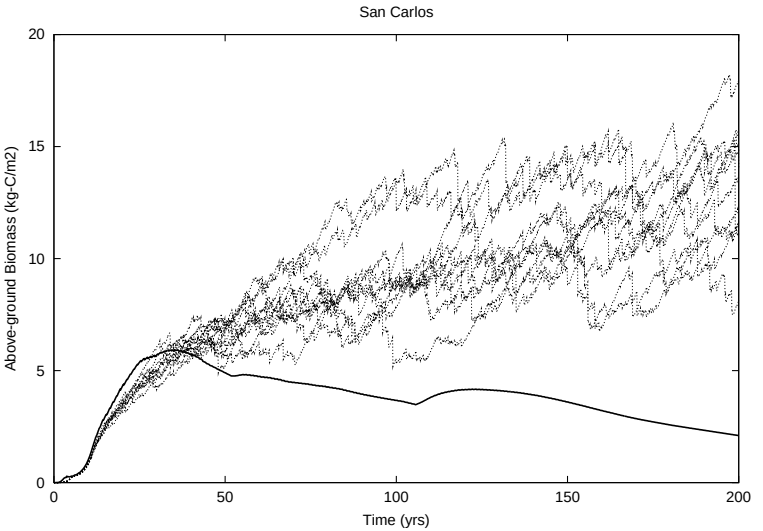

Fig. 5(a). 200 year trajectory in above-ground biomass (kgC m^{-2}) at the San Carlos tropical forest site (2°N, 68°W), predicted by the ED model implemented as a individual-based stochastic gap model. Figure shows 10 runs of the stochastic process with each run containing 25 gaps (dashed lines). Also shown (solid line) is the trajectory predicted by a traditional size-structured approximation (Eq. (1)) of the stochastic gap model.

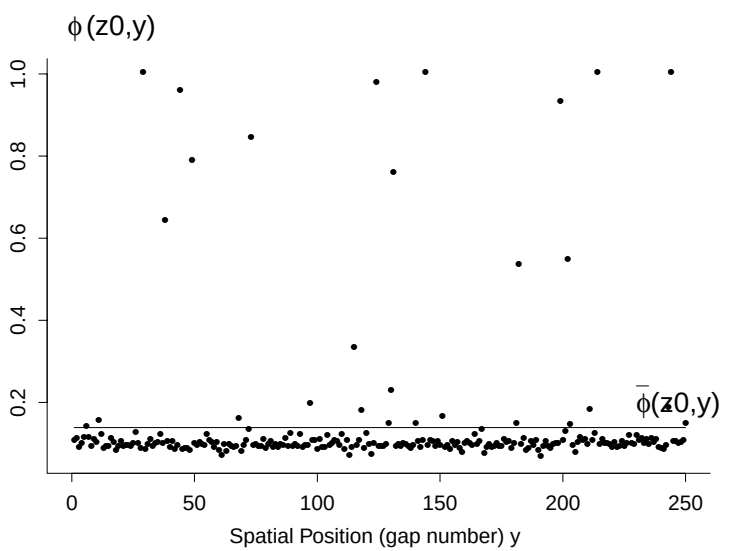

Fig. 5(b). Distribution of understorey light levels across gaps predicted by the model after the 200 year integration shown in (a). Horizontal line shows the average light level at the bottom of the plant canopy $\bar{\phi}(z_0, y)$ where size z_0 corresponds to a height of 0.5 m. This spatially-averaged (across-gap) light level is the effective understorey light level in the traditional size-structured approximation of the stochastic gap model shown in (a).

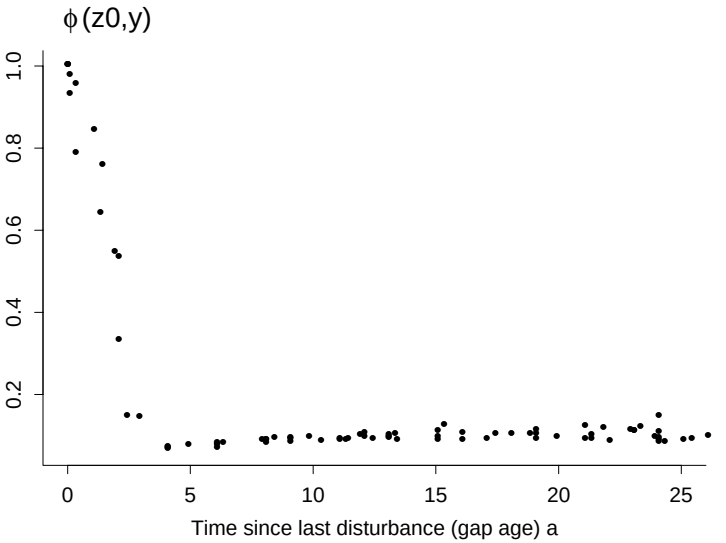

Fig. 5(c). Distribution of understorey light levels $\phi(\underline{z}_0, y)$ in gaps shown in (a) plotted as a function of their time since last disturbance a. As this figure shows, the time since disturbance a accounts for most of the spatial (across gap) variation in light-levels shown in (b).

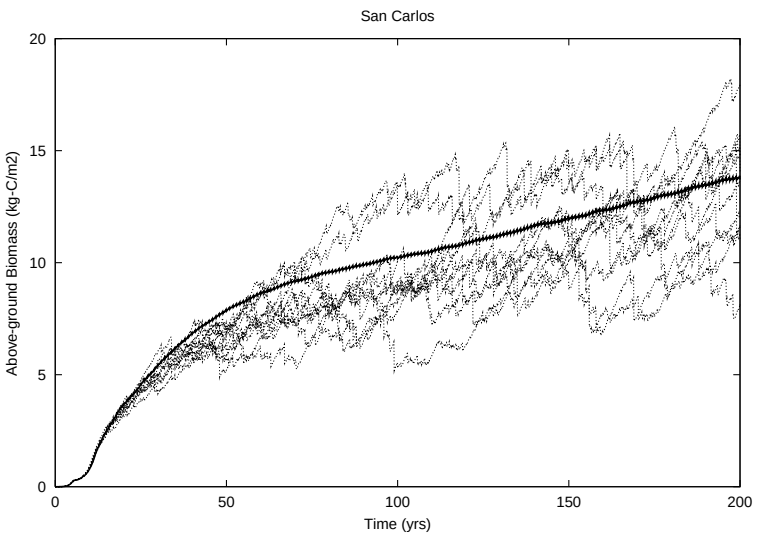

Fig. 5(d). Trajectory of above-ground biomass (kgC m^{-2}) at the San Carlos (SC) tropical forest site as output by 10 runs of ED implemented as a stochastic gap model (dashed lines) and the size-age structured (SAS) approximation (Eqs. (4) and(5)) (solid line).

Scaling

From a mathematical viewpoint, ED is a spatially-distributed, non-linear, stochastic process whose large-scale behavior we are interested in knowing. It is spatial because of local nature of competition among plants for light, water and nitrogen, non-linear due to the nature of the relationships between local resource availability and a plant's growth, mortality and recruitment rates, and stochastic because of the randomness of birth, dispersal, death and disturbance.

The key to scaling such a model is the recognition that the ensemble average used for stand or landscape-level predictions is, in the limit of a large number of runs, the first moment of the stochastic process. We seek to formulate a differential equation that captures time-dependent behavior of this first moment while taking account of the variability in the model behavior that arises due to the stochastic nature of the processes operating in the model. A standard mathematical technique is to formulate a partial differential equation (PDE) that accounts for the variability in the stochastic process by capturing the dynamics of those processes that give rise to important heterogeneities in the ecosystem (36–38).

A common initial strategy is to formulate a size-structured approximation that takes account of the size-related heterogeneity in light availability within the plant canopy

$$\underbrace{\frac{\partial}{\partial t}\left(n(\underline{z},\underline{x},t)\right)}_{\text{ch. in plant density}} = -\frac{\partial}{\partial \underline{z}_s}\underbrace{\left[g_s(\underline{z},\underline{x},\bar{r},t)n(\underline{z},\underline{x},t)\right]}_{\text{growth in stem}} - \frac{\partial}{\partial \underline{z}_a}\underbrace{\left[g_a(\underline{z},\underline{x},\bar{r},t)n(\underline{z},\underline{x},t)\right]}_{\text{growth in active tissues}} - \underbrace{\left[\mu(\underline{z},\underline{x},\bar{r},t)+\lambda_F(t)\right]n(\underline{z},\underline{x},t)}_{\text{mortality and disturbance}} \quad (1)$$

with the single boundary condition

$$\underbrace{n(\underline{z}_0,\underline{x},t) = \frac{\int_{\underline{z}_{0s}}^{\infty}\int_{\underline{z}_{0a}}^{\infty} n(\underline{z},\underline{x},t)f(\underline{z},\underline{x},\bar{r},t)d\underline{z}_a\underline{z}_s}{g_a(\underline{z}_0,\underline{x},\bar{r},t)+g_s(\underline{z}_0,\underline{x},\bar{r},t)}}_{\text{reproduction}} \quad (2)$$

and initial condition

$$\underbrace{n(\underline{z},\underline{x},0) = n_0(\underline{z},\underline{x})}_{\text{initial plant community}} \quad (3)$$

(39, 34). However, as we shall show, such size-structured approximations fail because in addition to the size-related vertical heterogeneity in light availability there is also a substantial degree of horizontal spatial heterogeneity in resource availability.

Note that in the size-structured approximation (Eq. (1)), $\bar{\underline{r}}$ is the spatially-averaged (non-local) resource conditions within the grid cell.

The traditional size-structured approximation fails because while it accounts for the vertical stratification in the light environment caused by plant shading, it does not account for the substantial degree of endogenous spatial heterogeneity in light availability between gaps (Fig. 5(b)). In the traditional size-structured approximation, the spatial variation in light profiles is averaged into a single mean light profile. For example, the line in Fig. 5(b) shows the light level assumed to be present in all gaps at a height of 0.5 m in the size-structured approximation. This light level is too low to allow regeneration to keep pace with canopy mortality, and so the size-structured approximation under-predicts the biomass in the simulator (solid line in Fig. 5(a)—see also 40).

One approach to overcoming this problem of spatial heterogeneity in resource availability is to develop a second-order approximation, that takes account of the covariance that develops between local resource availability $\underline{r}(\underline{z}, y, t)$ (here a vector of local light, water and nitrogen availability) and local plant density $n(\underline{z}, \underline{x}, y, t)$ (for example saplings of shade intolerant species are most abundant in locations with high light). This approach is often used in statistical physics where it is called "Gaussian Random Fields". However, although the second order approximations work for some simple models of plant competition (38, 41, 42, Kubo, pers. comm), they are not generally useful due to instabilities arising from the omission of higher order terms and the large number of covariance equations necessary for approximating a functionally diverse plant community.

An alternative approach, which we use in our SAS approximation, is to identify, characterize, and then condition upon, events that are responsible for generating horizontal spatial heterogeneity in resource availability. Two stochastic processes are responsible for generating most of the spatial heterogeneity in resource availability within the simulator. In tropical forest areas such as San Carlos, the majority of the spatial variation in light availability is associated with the mortality of large canopy trees. To see this, suppose we nominally define a canopy tree as any individual greater than 10 m in height. Suppose further that for each spatial position (gap) y within the gap simulator, we record the time since the last stochastic canopy tree death, and call this event a disturbance event, and the time since the last such event the age a of the gap. If we then plot light availability within each gap as function of its age, we see a clear pattern in the distribution of light environments (Fig. 5(c)), that is responsible for most of the between-gap scatter shown in Fig. 5(b). The second source of heterogeneity is fire which is responsible for most of the spatial heterogeneity resource availability in arid areas where they occur.

The SAS approximation takes account of both the horizontal spatial heterogeneity and size-related heterogeneity. It captures horizontal heterogeneity by keeping track of a, the time since the last death of large adult tree or disturbance event. More formally, we develop a size-structured approximation for the ensemble mean conditional on age a. The derivation of this conditional approximation is facilitated by a subtle change in the way we view the stochastic events within the model. Gap simulators work because the size of a modeled gap is similar to the size of a large tree's crown (43, 44). This ensures that single canopy tree deaths cause the high resource levels needed for rapid regeneration, and also causes gaps to contain at most one large tree. Suppose that we now replace the per-individual random coin tosses that cause mortality in the simulator, with a per-place coin toss with the same probability, but only for trees above a threshold size $\underline{z}^*$. That is, we continue to toss pseudo-random coins for each individual smaller than z^* with probability of mortality μ, and toss a single coin with the same probability for the entire gap. If this toss indicates mortality, then we kill every tree in the gap taller than z^*. Because there is typically at most one such tree (if z^* is 10–15 m or larger), this change has no effect on the predictions of the stochastic process, as simulations confirm. However, because large tree deaths are now exogenous place-centered events, we can treat canopy tree deaths like other place-centered disturbances in the model such as fires. As Fig. 5(c) suggests, by conditioning appropriately on the occurrence of disturbance events and keeping track of the changing distribution of ages a since disturbance, we can account for the horizontal and vertical heterogeneity in resource availability they introduce.

Size- and Age-Structured (SAS) approximation

In the SAS approximation, we use the new variable a to indicate the time since the most recent disturbance of any type (either fire or wind-throw). Note however in principle, a could be a vector, with each element denoting the time since the last event of a particular disturbance type. To write an expression for the change in plant density $n(\underline{z}, \underline{x}, a, t)$, we must redefine the transition probabilities in accordance with the new rules governing disturbance (i.e. the $\underline{z}^*$ assumption). $\mu(\underline{a}, \underline{x}, \underline{r}, t)\Delta t$ is now the probability of mortality from density-dependent causes and, for plants shorter than $\underline{z}^*$, from non-fire density-independent causes. Letting $\lambda_{DI}(a, y, t)\Delta t$ be the probability of a density-independent disturbance within a gap of age a that kills canopy trees taller than $\underline{z}^*$ and letting $\lambda(a, t) = \lambda_F(a, t) + \lambda_{DI}(a, t)$, where $\lambda_F(a, t)\Delta t$ is the probability of fire. Finally, we define $s(z, a, t)$ as a step function equal to $\lambda_{DI}(a, t)$ for $z < \underline{z}^*$, and $s(z, a, t) = 0$ otherwise.

$$\underbrace{\frac{\partial}{\partial t} n(\underline{z}, \underline{x}, a, t)}_{\text{ch. in plant density}} = -\frac{\partial}{\partial z_s}\underbrace{\left[g_s(\underline{z}, \underline{x}, \bar{\underline{r}}, t) n(\underline{z}, \underline{x}, a, t)\right]}_{\text{growth in stem}}$$

$$-\frac{\partial}{\partial z_a}\underbrace{\left[g_a(\underline{z}, \underline{x}, \bar{\underline{r}}, t) n(\underline{z}, \underline{x}, a, t)\right]}_{\text{growth in active tissues}}$$

$$-\underbrace{\frac{\partial}{\partial a} n(\underline{z}, \underline{x}, a, t)}_{\text{aging of plant community}} - \underbrace{\mu(\underline{z}, \underline{x}, \bar{\underline{r}}, t) n(\underline{z}, \underline{x}, a, t)}_{\text{mortality}}. \tag{4}$$

In addition, we require an equation for the probability distribution of age states a. Let $p(a, t)$ be the distribution of times since disturbance. Recalling that disturbances are occurring at rate $\lambda(a, t) = \lambda_F(a, t) + \lambda_{DI}(a, t)$, we can derive a PDE for the dynamics of $p(a, t)$, the age structure of the ecosystem within the grid cell.

$$\underbrace{\frac{\partial}{\partial t} p(a,t)}_{\text{ch. in age structure}} = \underbrace{-\frac{\partial}{\partial a} p(a,t)}_{\text{aging}} - \underbrace{\lambda(a,t) p(a,t)}_{\text{disturbance}}, \tag{5}$$

and

$$\int_0^\infty p(a,t)da = 1.$$

The first term describes the aging process while the second $\lambda(a, t)$ term gives the rate at which areas of age a are disturbed. Equation (5) is the Von-Foerster age distribution equation (45).

For the case of random dispersal between gaps within a grid cell the recruitment of new seedlings $f(\underline{z}, \underline{x}, a, t)$ corresponds to a flux of individuals into the system at $(\underline{z}_0, a)$ giving the following Neumann boundary condition

$$\underbrace{n(\underline{z}_0, \underline{x}, a, t) = \frac{\int_0^\infty \int_{\underline{z}_{0s}}^\infty \int_{\underline{z}_{0a}}^\infty d(\underline{z}, \underline{x}, a, t) f(\underline{z}, \underline{x}, a, t) p(a,t) d\underline{z}_A d\underline{z}_S da}{g_a(\underline{z}_0, \underline{x}, \bar{\underline{r}}, t) + g_s(\underline{z}_0, \underline{x}, \bar{\underline{r}}, t)}}_{\text{recruitment}}. \tag{6}$$

Equation (4) also has a second boundary condition, describing the state of the ecosystem following a disturbance event:

$$\underbrace{n(\underline{z}, \underline{x}, 0, t) = \int_0^\infty s\big(h(\underline{z}, \underline{x})a, t\big)n(\underline{z}, \underline{x}, a, t)da.}_{\text{plant community following disturbance event}} \tag{7}$$

The fraction of newly disturbed areas $p(0, t)$ is given by the boundary condition

$$\underbrace{p(0,t) = \int_0^\infty \lambda(a,t)p(a,t)da.}_{\text{formation of newly disturbed areas}} \tag{8}$$

We complete the SAS approximation by specifying initial conditions for (5) and (4) corresponding respectively to the initial age distribution of areas within the grid cell and the size distribution of the plant types within each of these areas

$$\underbrace{n(\underline{z}, \underline{x}, a, 0) = n_0(\underline{z}, \underline{x}, a)}_{\text{initial plant community}}, \tag{9}$$

$$\underbrace{p(a, 0) = p_0(a)}_{\text{initial age distribution}}. \tag{10}$$

The PDEs (Eqs. (4) and (5)) and their associated boundary (6)–(8), and initial conditions (9) and (10) describe the dynamics of a SAS plant community within a grid cell Ω, where $n(\underline{z}, \underline{x}, a, t)$ is formally the expected density of plants of size $\underline{z}$ and type $\underline{x}$ in a gap of age a at time t.

Following precisely the same steps used to derive the PDEs above, we also derive equations for the below-ground water, carbon and nitrogen conditional on age a that describe the below-ground dynamics of the ecosystem as function age.

SAS PDEs similar to (4) and (5) have been used previously to model forest dynamics (46, 47), however in these studies, the equations were formulated at the stand-level rather than as an approximation to an individual-based model. The stand-level equations in these studies would approximate individual-based models if the appropriate changes were made to the mortality functions (the $\underline{z}^*$ assumption).

As Fig. 5(d) shows, unlike the traditional size-structured approximation, the SAS approximation captures the ensemble means across a wide range of conditions. Note that the SAS approximation of total above-ground carbon (red) predicts the center of the ensemble of stochastic runs (green) in all climates from very dry to very wet, and from strongly to weakly seasonal. In subsequent sections, we show that the SAS approximation also accurately predicts the ensemble means of the biomass of each functional type.

Regional scale implementation

Using the SAS approximation, is to possible to implement ED at regional scales. Figure 6 shows the pattern of above-ground biomass obtained from a 200

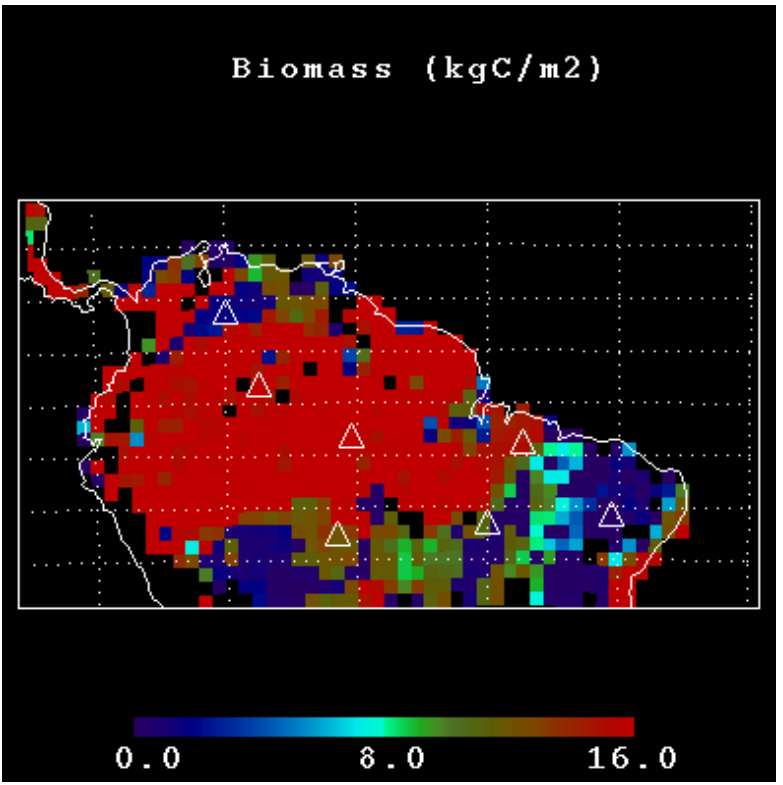

Fig. 6. Model predictions of average biomass compared after 200 year regional model integration using the SAS PDEs.

year integration of the ED model for tropical and sub-tropical South America (15°N to 15°S latitude). In the figure one can see the low biomass grassland region of the Orinoco Llanos, the large forested region of Amazonia with high biomass, the Atlantic coastal tropical forest, the comparatively low biomass region of the dry Caatinga in the South East, and the deserts along the Pacific coast.

San Carlos: patterns of forest succession

San Carlos is the evergreen rainforest site at which the chronosequence in of above-ground biomass shown in Fig. 3 was collected. Does our biosphere model correctly predict the observed long timescale of carbon accumulation at San Carlos? If so, then what mechanisms cause the delay? And are these the same mechanisms that occur in nature?

The pattern of accumulation exhibited our model closely resembles the observed trajectory of above-ground biomass recorded in the two-hundred year chronosequence at San Carlos (Fig. 7(a)). Biomass accumulation is initially

rapid, with approximately 6–8 kg-C m^{-2} of accumulation during the first 30–50 years. Then after this initial period of rapid increase, above-ground biomass accumulates more slowly, gaining a further 6–7 kg-C m^{-2} over the next 150–170 years and reaching 13 kg-C m^{-2} after 200 years (Fig. 7(a)).

The mechanism responsible for this 200 year timescale and pattern of biomass accumulation is the same in the model and observations. As Saldarriaga *et al.* (27) note in their paper, the rapid initial biomass increase during the first 50 years is caused by colonization by fast-growing early successional trees with low wood density that rapidly form a closed forest canopy. Above-ground biomass then continues to accumulate for the next 150 years, albeit more slowly, due to gradual replacement of the early successional trees with slower growing, mid and late successional trees with higher wood density.

The pattern and timescale of above-ground biomass succession at San Carlos predicted by our model arises from a similar successional processes (Fig. 7(b)). After a short period in which grasses briefly proliferate, early successional tree growth dominates biomass accumulation during the first 50 years. This is followed by slower biomass accumulation as the early successional trees are competitively replaced by slower growing mid and late successional trees with higher wood densities. Note also that the SAS approximation shown in Fig. 7(b) accurately predicts the ensemble average for each functional type obtained from stochastic simulations of ED. These runs are the same as those used to produce the pattern of total biomass in Fig. 5(d).

In summary, two processes produce the long timescale of succession at San Carlos in both the observations and our model. First, height-structured competition allows fast growing but short-lived trees to forestall domination of the forest by

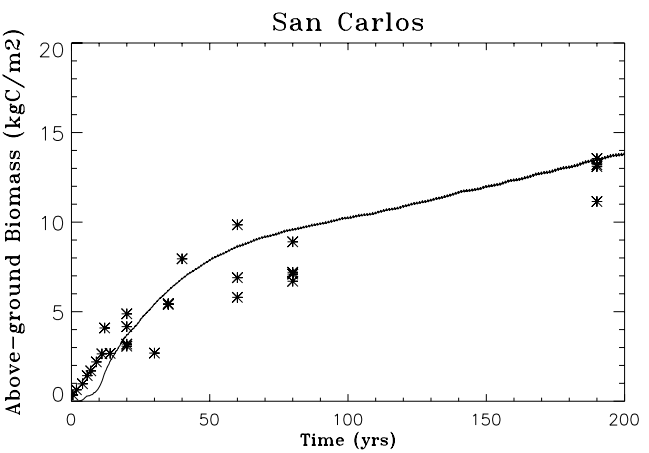

Fig. 7(a). Trajectory of above-ground biomass (kg-C m^{-2}) at San Carlos predicted by the ED model compared to the chronosequence of above-ground biomass measurements made by Uhl (28), and Saldarriaga *et al.* (27).

higher wood density species for a century or so. These high wood density species are relatively long-lived and thus grow to large average size and store large amounts of carbon. Second, a further century or so is required before the average size of late successional trees stops increasing. Thus, the first mechanism is at the community level (temporary competitive suppression of the eventual dominant), while the second is at the population level (the time required for the formation of a stable size and age distribution of the late successional dominant). Also, it is important to understand that sub-grid scale heterogeneity is essential to matching the predicted and actual patterns at San Carlos. The traditional size-structured approximation, which lacks horizontal heterogeneity caused by sub-grid scale disturbance, severely under-predicts the trajectory of above-ground biomass at San Carlos (compare Figs. 3 and 5(a)).

The predictions of ED illustrate the advantages of formulating ecosystem models at the scale of individual plants. Individual-based ecosystem models naturally capture the fine-scale population and community-level processes responsible for the slow timescale of carbon uptake in aggrading tropical forest

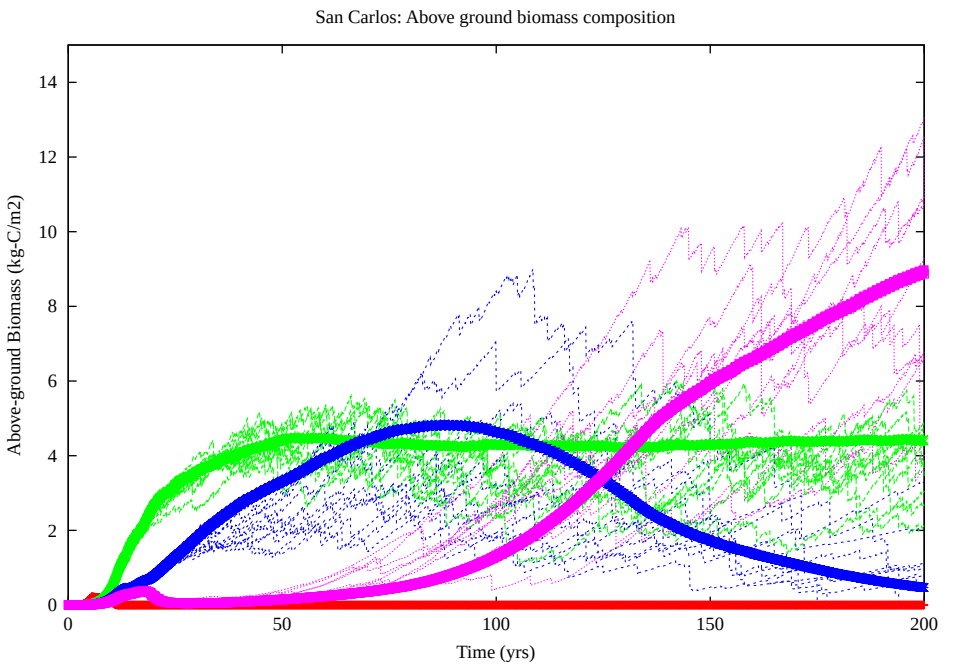

Fig. 7(b). Forest composition changes at the San Carlos tropical forest site underlying the trajectory of total above-ground biomass shown in (a). Thin lines show the trajectories of above-ground biomass of C_4 grass (red), early successional trees (green), mid-successional trees (blue) and late successional trees (pink) predicted by the ED model implemented as an individual-based stochastic gap simulator. Thick lines show the corresponding trajectories from a SAS approximation.

at San Carlos These and related factors explain the continued successes of individual-based ecosystem models over the past 30 years (48, 49, 33). Also, because the model is formulated at a scale consistent with field studies the model is comparatively easier to parameterize and test with data collected at finer scales (48, 34). For example, abundant data on allometry of individual trees can be used to constrain ecosystem-level allocation and forest inventory data can be used to test the model's predictions about forest composition, stand structure and carbon storage. In addition, many forms of satellite data are now approaching resolutions consistent with formulating fine-scale models of the biosphere.

Compared to the traditional kind of ecosystem models used to produce large scale predictions of ecosystem properties that are parameterized and tested at single spatial scale, fine-scale formulations are able to connect to data collected at variety scales is also likely to improve confidence in large-scale predictions, which are often difficult to test directly. For example, our model predicts many of the details of ecosystem dynamics at San Carlos including successional transients, and has regional predictions of NPP and above-ground carbon and soil carbon that are generally consistent with the limited regional data available.

The most general contribution of our study is the scaling methods that provide the PDEs for predicting ensembles of gap-model runs. We suspect that these will work for any individual-based simulator of vegetation. The physical environmental sciences rely fundamentally on a directly analogous scaling technologies. Modern atmospheric and oceanic GCMs typically rely on PDEs rather than stochastic Lagrangian particle simulators for reasons of computational efficiency and because their compactness permits mathematical analysis. The scaling that leads from the individual-based gap model to our SAS approximation is analogous to that leading from a stochastic Lagrangian particle simulator to the Navier-Stokes equations.

The fundamental difference between the Navier-Stokes equations and the SAS PDEs is that rules governing the underlying stochastic process are completely understood for the former, but still very much under development for the latter. We emphasize that the individual-based model introduced in this paper is only one of many possible formulations for a stochastic gap model driven by ecophysiological and biogeochemical mechanisms.

The success of the SAS approximation in capturing the dynamics of the corresponding stochastic gap model implies that ecosystem dynamics at the grid-cell scale depend critically on the size structure and the local disturbance history (age distribution) at sub-grid scales. This point is further emphasized by the success of the SAS approximation at San Carlos, where field data confirm the critical role of sub-grid scale processes in large-scale ecosystem structure and dynamics. It also confirms the findings of earlier work using stochastic gap models which showed the importance of disturbances such as fire and wind-throw in influencing the structure and composition of plant communities (50–52).

Since the information regarding the long term behavior of stochastic gap model resides in the sub-grid cell heterogeneity of the SAS approximation, we can view the three components of sub-grid cell variability: size structure, age

structure and functional type composition, as embodying the long-term memory of the above-ground ecosystem which determines its subsequent dynamics. As a result, the ability to characterize and understand the size-related, age-related, and composition-related structure of ecosystems is fundamental to a better understanding of their long-term fate. This further suggests that inventories of vegetation structure (e.g. Phillips *et al*. (53)), may hold the key to predicting the future large-scale dynamics of ecosystems, and provides additional incentive for satellite measurements of vegetation structure (Dubayah *et al*. (54)).

Use of the SAS approximation offers the same two principal advantages over direct simulation of stochastic gap models as the Navier-Stokes equations offer over Lagrangian particle simulators. First, the formulation obviates the need for many-run stochastic ensembles, thereby greatly reducing the computational burden (by one to several orders of magnitude). Accurate characterization of ensemble means through simulation can be a computationally intensive exercise. Our experience suggests that the necessary number is often one to several orders of magnitude larger. For this reason, we found it necessary to simulate at least 250 gaps to characterize ensemble means in this paper (10 runs each with 25 gaps). If the size of the ensembles were too small, then one or more of the functional types that coexist in our simulations of evergreen rainforest at San Carlos would become extinct with high probability, and this would significantly alter the timescale of carbon accumulation shown in Fig. 7(a). The necessary size of an ensemble increases geometrically with the number of dimensions of heterogeneity that must be captured. Future studies may find it necessary to move beyond our formulation and include sub-grid scale heterogeneity in land-use, abiotic physical heterogeneities such as soils and topography, as well as other dimensions of functional diversity. Such applications will amplify the importance of the increase in computational efficiency offered by the SAS PDEs.

Second, like the Navier-Stokes equations, the SAS approximation offers the promise of increased understanding through mathematical analysis. Although analytically formidable, the SAS approximation is much more tractable than the stochastic gap model itself. We suspect that it will ultimately yield analytical insights about the connection between local ecosystem processes and large-scale ecosystem function.

Predicting the fate of the U.S. carbon sink

Hurtt *et al*. (in review) implemented the ED model for the coterminous U.S. to predict the future of the U.S. carbon sink. In this study, ED was implemented on a 1 degree grid and run for 700 years (1400–2100). Information about soils in the coterminous U.S. was taken from (55, 56). We did not include interannual changes in climate (hourly weather in each 1 grid cell was interpolated from the ISLSCP climatology (55, 56) which reflects the recent climate) and used a constant modern concentration of atmospheric CO_2 (350 ppmv). If the conclusions of the recent VMAP study (10) are correct, then omission of climate change and increasing atmospheric CO_2 should have caused ED to underestimate the U.S.

carbon sink. However, because the analysis of FIA data in Study 2 (above) implies that effects of climate change and CO_2 fertilization have been negligible to date, we chose the conservative course of omitting these effects.

The model was first spooled-up for 300 years to produce the potential vegetation of the continent. In simulated year 1700, the model predicted a distribution of biomes, net primary productivity, and carbon storage in generally close agreement with information from empirical studies, except in the far Pacific northwest where our climatology was too crude to resolve the coastal climate (14). Also, approximately 80 million hectares of land burned annually, which compares favorably with national fire statistics (6, 57) for the period prior to 1840.

During simulated years 1700–2000, a reconstruction of the historical pattern of land use was imposed, giving the amount of land in each year up in six categories: cropland, pasture, primary forest, secondary forest, softwood plantation, and unmanaged non-forest. The reconstruction was very similar to the one in (6), because it was derived in the same way from census, land use, forest and fire statistics. The minor differences were caused by our reliance on USFS data for the amount of forest land after 1952 rather than USDA agricultural census data (58). We made this substitution because the USDA census classifies many forested lands (i.e. parks, military bases) as "Other". In addition, we used a 1 map of potential vegetation (59) and a 0.5 reconstruction of U.S. croplands (59) to increase the spatial resolution of the land use reconstruction in (6) from seven homogenous regions to 1. Finally, we used a new maximum likelihood method, together with the FIA data, to estimate, for each of five regions, harvest rates for secondary forests and plantations as functions of the time since last harvest. Both harvest rates of primary forest and the national pattern of fire suppression were as in (6). Harvest practices were taken from (60), including the fractions removed from the site and left as slash.

In summary, the historical pattern of land use in Hurtt *et al.* (in review) begins with potential vegetation in 1700. In each subsequent year, the category totals for the seven regions considered in (6) match the values derived from censuses and other data, the spatial pattern matches, at 1, the croplands reconstruction in (59), and the extent of fire matches the wildfire record (14). Also, at the end of the reconstruction, modeled forest stand age distributions match the distributions given by the FIA data base (14). The reconstruction's 1990 map compares favorably to a recent land classification derived from satellite imagery (14).

From 2000–2100, the model assumes that current land use practices will continue. In the best-case scenario, fire suppression is also assumed to remain effective, despite increases in fuel wood. The study forecasts the fate of only that portion of the U.S. sink caused by vegetation on non-croplands. The model predicts that this part of the sink will decrease from approximately 0.4 PgC y^{-1} in 1990, to 0.2 PgC y^{-1} in 2050 and 0.1 PgC y^{-1} in 2100. Approximately half this sink is caused by fire suppression, while the other half is caused by recovery of forests following the clear cutting of old growth in the 19th and 20th centuries.

The reason for the expected decrease in the sink is simply that U.S. ecosystems begin to equilibrate during the 21st century to current patterns of harvesting and fire. In contrast, a sink caused by CO_2 or nitrogen fertilization or climate change might, is some cases, be expected to increase throughout the coming century.

CONCLUSIONS

1. The carbon sink in the coterminous U.S. is large. It stores annually from one third to two thirds of a billion metric tons of carbon.

2. The sink is caused by a variety of factors; but land use and management cause the overwhelming majority of it.

3. The land use sink will decrease throughout the current century, as U.S. ecosystems adjust to management and land use.

REFERENCES

(1) P. P. Tans, I. Y. Fung and T. Takahashi, *Science* **247**, 1531 (1990); M. Battle, M. Bender, T. Sowers, P. P. Tans, J. H. Butler, J. W. Elkins, J. T. Conway, N. Zhang, P. Lang and A. D. Clark, *Nature* **383**, 2312 (1996); P. Ciais, P. P. Tans, M. Trolier, J. W. C. White and R. J. Francey, *Science* **269**, 1098 (1995); R. F. Keeling, S. C. Piper and M. Heimann, *Nature* **381**, 218 (1996); J. T. Randerson, M. V. Thompson, T. J. Conway, I. Y. Fung and C. B. Field, *Global Biogeochem. Cycles* **11**, 535 (1997); D. Schimel, I. G. Enting, M. Heimann, T. M. L. Wigley, D. Raynaud, D. Alves and U. Siegenthaler, *Climate Change 1995*, J. T. Houghton, L. G. M. Filho, B. A. Callander, N. Harris, A. Kattenberg and K. Maskell (eds.), Cambridge University Press, Cambridge, 25–71 (1995).

(2) S. Fan, M. Gloor, J. Mahlman, S. Pacala, J. Sarmieto, T. Takahashi and P. Tans, *Science* **282**, 442 (1998).

(3) P. J. Rayner, I. G. Enting, R. J. Francey and R. Langenfields, *Tellus* **51B**, 213 (1999); T. Kaminski, M. Heimann and R. Giering, *J. Geophys. Res*. **104**, 18555 (1999); P. Bousquet, P. Ciais, P. Peylin, M. Ramonet and P. Monfray, *J. Geophys. Res*. **104**, 26161 (1999); P. Bousquet, P. Ciais, P. Peylin, M. Ramonet and P. Monfray, *J. Geophys. Res*. **104**, 26179 (1999); P. Bousquet, P. Ciais, P. Peylin and P. Monfray, *Geophysical Monograph* **114** (1999).

(4) R. A. Houghton, J. L. Hackler and K. T. Lawrence, *Science* **285**, 574 (1999); R. A. Houghton and J. L. Hackler, *Global Ecology and Biogeography* **9**, 125 (2000); R. A. Houghton and J. L. Kackler, *Global Ecology and Biogeography* **9**, 145 (2000); D. P. Turner, G. J. Koerper, M. E. Harmon and J. J. Lee, *Ecological Applications* **5**, 421 (1995); R. A. Birdsey, Carbon Storage and Accumulation in the United States Forest Ecosystems, Forest Service General Technical Report WO-59, U.S. Dept. of Agriculture, Washington, D.C. (1992); R. A. Birdsey and L. S. Heath, *Productivity of America's Forest Ecosystems*, L. A. Joyce (ed.), Forest Service General Technical Report RM-GTR-271, Fort Collins, CO, 56–70 (1995); S. L. Brown and P. E. Schroeder, *Ecological Applications* **9**, 968 (1999).

(5) R. K. Dixon, S. Brown, R. A. Houghton, A. M. Solomon, M. C. Trexler and J. Wisniewski, *Science* **263**, 185 (1994); P. E. Kauppi, K. Mielikaiinen and K. Kuusela, *Science* **256**, 20 (1997); G. J. Nabuurs, G. J. Pavienen, R. Sikkema and G. M. J. Mohren, *Biomass Bioenergy* **13**, 345 (1997); R. A. Houghton, J. L. Hackler and K. T. Lawrence, *Science* **285**, 547 (1999); W. A. Kurz and M. J. Apps, *Ecological Applications* **9**, 526 (1999); J. L. Sarmiento and S. Wofsy (eds.), *A U.S. Carbon Cycle Science Plan*, University Corporation for Atmospheric Research, Boulder, CO (1999).

(6) R. A. Houghton, J. L. Hackler and K. T. Lawrence, *Science* **285**, 574 (1999); R. A. Houghton and J. L. Kackler, *Global Ecology and Biogeography* **9**, 125 (2000); R. A. Houghton and J. L. Kackler, *Global Ecology and Biogeography* **9**, 145 (2000).

(7) R. A. Birdsey, Carbon Storage and Accumulation in the United States Forest Ecosystems, Forest Service General Technical Report WO-59, U.S. Dept. of Agriculture, Washington, D.C. (1992); R. A. Birdsey and L. S. Heath, *Productivity of America's Forest Ecosystems*, L. A. Joyce (ed.), Forest Service General Technical Report RM-GTR-271, Fort Collins, CO, 56–70 (1995); D. S. Powell, J. L Faulkner, D. R. Darr, Zhiliang Zhu and D. W. MacCleery, Forest Resources of the United States, 1992, Forest Service General Technical Report RM-234-Revised (1994).
(8) D. P. Turner, G. J. Koerper, M. E. Harmon and J. J. Lee, *Ecological Applications* **5**, 421 (1995).
(9) R. F. Stallard, *Global Biogeochem. Cycles* **12**, 231 (1998).
(10) D. Schimel, J. Melillo, A. D. Mcguire, D. Kicklighter, T. Kittel, N. Rosenbloom, S. Running, P. Thornton, D. Ojima, W. Parton, R. Kelly, M. Sykes, R. Neilson, B. Rizzo and L. Pitelka, *Science* **287**, 2004 (1999).
(11) S. W. Pacala, G. C. Hurtt, D. Baker, P. Peylin, R. A. Houghton, R. A. Birdsey, L. Heath, E. T. Sundquist, R. F. Stallard, P. Ciais, P. Moorcroft, J. Caspersen, E. Shevliakova, B. Moore, G. Kohlmaier, E. Holland, M. Gloor, M. E. Harmon, S.-M. Fan, J. L. Sarmiento, C. Goodale, D. Schimel and C. B. Field, *Science*, June 22 (2001).
(12) J. P. Caspersen, S. W. Pacala, J. C. Jenkins, G. C. Hurtt, P. R. Moorcroft and R. A. Birdsey, *Science* **290**, 1148 (2001).
(13) P. R. Moorcroft, G. C. Hurtt and S. W. Pacala, *Ecological Monographs* (in press).
(14) G. C. Hurtt, S. W. Pacala, P. Moorcroft, J. Caspersen and E. Shevliakova (in review).
(15) W. M. Post and K. C. Kwon, *Global Change Biology* **6**, 317 (2000); D. W. Johnson, *Water, Air and Soil Pollution* **64**, 83 (1992).
(16) Woody detritus database in: M. E. Harmon, J. M. Harmon, O. N. Krankina and M. Uatskov, A global database for stores and dynamics of woody detritus (2000).
(17) K. E. Skog and G. A. Nicholson, *Forest Products Journal* **48**, 75 (1998).
(18) A. S. Donigian, A. S. Patwardin, R. B. Jackson, T. O. Barnwell, K. B. Weinrich and A. L. Rowell, *Modeling Impacts of Agricultural Practices on Soil Carbon in the Central U.S.*, R. Lal *et al.* (eds.). CRC Press, 121 (1995); R. Q. Cannell and J. D. Hawes, *Soil Tillage Res.* **30**, 245 (1994); D. L. Gebhart, H. B. Johnson, H. S. Mayeaux and H. W. Polley, *J. of Soil and Water Cons.* **49**, 488 (1994); P. Smith, D. S. Powlson, J. J. Glendining and J. U. Smith, *Global Change Biology* **4**, 679 (1998).
(19) Food and Agriculture Organization (FAO), FAO Production Yearbook 1990; FAO, FAO Production Yearbook 2000.
(20) U.S. Department of Agriculture (USDA), Agricultural Statistics 2000, USDA Nat. Agr. Stat. Service, Washington, D.C. (2000), http://www.usda.gov.nass.pubs/agr00/acro00.htm.
(21) E. T. Sundquist, R. F. Stallard and N. Bliss (in preparation).
(22) The FIA database is publicly available at: www.srsfia.usfs.msstate.edu/scripts/ew.htm.
(23) M. H. Hansen, T. Frieswyk, J. F. Glover and J. F. Kelly, General Technical Report NC-151 (USDA Forest Service 1992).
(24) R. A. Birdsey and H. T. Schreuder, General Technical Report RM-214 (USDA Forest Service 1992).
(25) To quantify current vital rates we selected plots that may have been selectively harvested but were not clearcut or otherwise heavily disturbed after the first inventory. We excluded clearcut or heavily disturbed plots by eliminating plots whose age was reset in the second inventory. In addition, we excluded plots occurring in plantations and plots that were classified as non-forest in either the first or the second inventory. We specify growth with an asymptotic Michaelis-Menten function: $G(A) = c + d(1 - \exp(-(e/d)A)) + \varepsilon$, where $\varepsilon \sim N(0, i + jG)$. We obtained maximum likelihood estimates of the growth and mortality parameters using a simulated annealing algorithm as we do in all subsequent analyses.
(26) Given $G(A)$ and μ we can calculate $B(A)$ for any value of $B(0)$. We estimate $B(0)$ as the value that provides the best fit to $B(A)_{obs}$. Here, we assume that $B(A)_{obs}$ is normally distributed with constant variance.
(27) J. Saldarriaga, D. C. West, M. L. Tharp and C. Uhl, Long-term chronosequence of forest succession in the upper Rio Negro of Colombia and Venezuela. *Journal of Ecology* **76** (1988).

(28)C. Uhl, Factors controlling succession following slash-and-burn agriculture in Amazonia. *Journal of Ecology* **75** (1987).
(29)C. Uhl and C. Jordan, Succession and nutrient dynamics following forest cutting and burning in Amazonia. *Ecology* **65** (1984).
(30)J. W. Reich, E. B. Rastetter, J. M. Melillo, D. W. Kicklighter, P. A. Steudler, B. J. Peterson, A. L. Grace, B. Moore, III and C. J. Vorosmarty, Potential net primary productivity in South America: application of a global model. *Ecological Applications* **1**(4) (1991).
(31)C. S. Potter, J. T. Randerson, C. B. Field, P. A. Matson, P. M. Vitousek, H. A. Mooney and S. A. Klooster, Terrestrial ecosystem production: a process model based on global satellite and surface data. *Global Biogeochemical Cycles* **7**(4) (1993).
(32)J. A. Foley, I. C. Prentice, N. Ramankutty, S. Levis, D. Pollard, S. Sitch and A. Haxeltine, An integrated biosphere model of land surface processes, terrestrial carbon balance, and vegetation dynamics. *Global Biogeochemical Cycles* **10**(4) (1996).
(33)M. Huston, Individual-based forest succession models and the theory of plant competition. In *Individual-Based Models and Approaches in Ecology*, D. DeAngelis and L. Gross (eds.), Routledge, Chapman and Hall, London (1992).
(34)S. W. Pacala, C. D. Canham, J. Saponara, J. A. Silander, R. K. Kobe and E. Ribbens, Forest models defined by field measurements: estimation, error analysis and dynamics. *Ecological Monographs* **66**(1) (1996).
(35)P. R. Moorcroft, G. C. Hurtt and S. W. Pacala, *Ecological Monographs* (in press).
(36)N. T. Bailey, *The Elements of Stochastic Processes with Applications to the Natural Sciences*, Wiley, New York (1964).
(37)J. Murray, *Mathematical Biology*. Springer-Verlag, Berlin (1990).
(38)S. Levin and S. Pacala, Theories of simplification and scaling in ecological systems. In *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, D. Tilman and P. Kareiva (eds.), Princeton University Press, Princeton, NJ (1998).
(39)A. Okubo, Diffusion and ecological problems: mathematical models, *Volume 10 of Biomathematics*. Springer-Verlag, Basel, Germany (1980).
(40)S. W. Pacala and D. H. Deutschman, Details that matter: the spatial distribution of individual trees maintains forest ecosystem function. *Oikos* **74** (1997).
(41)B. M. Bolker and S. W. Pacala, Using moment equations to understand stochastically driven spatial pattern formation in ecological systems. *Theoretical Population Biology* **52** (1997).
(42)S. Pacala and S. Levin. Biologically generated spatial pattern and the coexistence of competing species. In *Spatial Ecology: The Role of Space in Population Dynamics and Interspecific Interactions*, D. Tilman and P. Kareiva (eds.), Princeton University Press, Princeton, NJ (1998).
(43)D. Botkin, J. F. Janak and J. R. Wallis, Some ecological consequences of a computer model of plant growth. *Ecology* **60** (1972).
(44)H. H. Shugart and D. C. West, Development of an Appalacian deciduous forest model and its application to assessment of the impact of the chestnut blight. *Journal of Environmental Management* **5** (1977).
(45)von Foerster, H., *The Kinetics of Cellular Proliferation*, F. Stohlman, Jr. (ed.), Grune and Stratton, New York (1959).
(46)T. Kohyama, Size structured tree populations in gap-dynamic forest—the forest architecture hypothesis for the stable coexistence of species. *Journal of Ecology* **81** (1993).
(47)T. Kohyama and N. Shigesada, A size-distribution-based model of forest dynamics along a latitudinal environmental gradient. *Vegetatio* **121** (1995).
(48)H. H. Shugart and T. M. Smith, A review of forest patch models and their application to global change research. *Climate Change* **34** (1996).
(49)M. Huston, D. DeAngelis and W. Post, New computer models unify ecological theory. *Bioscience* **38** (1988).
(50)I. Noble and R. Slatyer, Concepts and models of succession in vascular plant communities subject to recurrent fire. In *Fire and the Australian Biota*, A. Gill, R. Groves and I. Noble (eds.), Australian Academy of Science, Canberra (1981).

(51)T. Doyle, The role of disturbance in the gap dynamics of a montane rain forest: an application of a tropical forest succession model. In *Forest Succession*, Springer-Verlag, Berlin (1981).
(52)H. Shugart and S. Seagle, Modeling forest landscapes and the role of disturbance in ecosystems and communities. In *Natural Disturbance: An Ecological and Evolutionary Perspective*, S. Pickett and P. White (eds.), Academic Press (1984).
(53)O. Phillips, Y. Malhi, N. Higuchi, W. Laurance, P. Nunez, R. Vasquez, S. Laurance, L. Ferreira, M. Stern, S. Brown and J. Grace, Changes in the carbon balance of tropical forests: Evidence from long-term plots. *Science* **282** (1998).
(54)R. Dubayah *et al.*, The vegetation canopy lidar mission. In *Proceedings of Land Satellite Information in the Next Decade*, Volume II, ASPRS (1997).
(55)B. W. Meeson, F. E. Corprew, J. M. P. McManus, D. M. Myers, J. W. Closs, K. J. Sun, D. J. Sunday and P. J. Sellars, ISLSCP Initiative 1 degree global data sets for land-atmosphere models, 1987–1988, CD Tech Report NASA (1995).
(56)P. J. Sellers, B. W. Meeson, J. Closs, J. Collatz, F. Corprew, D. Dazlich, F. G. Hall, Y. Kerr, R. Koster, S. Los, K. Mitchell, J. McManus, D. Myers, K. J. Sun and P. Try, An overview of the ISLSCP Initiative I global data sets. ISLSCP Inititative 1—global datasets for land—atmosphere models, 1987–1988, CD, Tech Report NASA (1995).
(57)R. A. Houghton and J. L. Hackler, *Global Ecology and Biogeography* **9**, 145 (2000); S. J. Pyne, *America's Fires. Management of Wildlands and Forests, Forest Historical Society*, Durham, North Carolina (1997); U.S. Department of Agriculture Forest Service, Cooperative Fire Protection Program 1926–1990, U.S. Government Printing Office, Washington, D.C. (1994).
(58)U.S. Department of Agriculture, Major Land uses (1945–1992), Stock #89003, Economic Research Service, USDA, Washington, D.C. (1996).
(59)N. Ramankutty and J. A. Foley, *Global Biogeochemical Cycles* **13**, 4 (1999).
(60)M. E. Harmon and J. J. Lee, *Ecological Applications* **5**, 421 (1995).

S. W. Pacala (e-mail: steve@eno.princeton.edu), G. C. Hurtt, P. R. Moorcroft and J. P. Caspersen