12 SYMFOR Technical Note Series

An ecological model for the management of natural forests in the Tapajos region, Amazonian Brazil

Paul D. Phillips
Celso P. de Azevedo
Ian S. Thompson
José N.M. Silva

2002

Centre for the Study of Environmental Change and Sustainability (CECS)
The University of Edinburgh
John Muir Building
Mayfield Road
Edinburgh EH9 3JK
United Kingdom
Tel: +44 131 650 7860
Fax: +44 131 650 7863
Email: info@symfor.org
http://www.symfor.org/technical

DFID Department for International Development This publication is an output from a project funded through the Forestry Research Programme of the UK Department for International Development (DFID) for the benefit of developing countries. The views expressed are not necessarily those of DFID. R6915 Forestry Research Programme.

Summary

An ecological model describing the processes of growth, recruitment and mortality was developed using the EMBRAPA SFC dataset from Amazonian Brazil. The model was implemented in the SYMFOR framework and calibrated and tuned with simulations of primary forest dynamics. The model was tested by comparing data with simulations of logged-over forest dynamics.

This document describes the data and development of the model, including species groups and the sub models for growth, recruitment and mortality. These were integrated within the SYMFOR framework for evaluation and application. Results of typical simulations are presented as to enable assessment of the performance of the model.

This document is intended to document the process of model development in such a way that would facilitate the development of similar models for other areas, forest types or applications.

The full text of this document is available from http://www.symfor.org/technical.

Contents

1 Introduction

Sustainable forest management requires knowledge of the impacts of management on future forest structure and composition. Acquisition of such knowledge often requires the application of a detailed, flexible modelling system calibrated for the local region. Prediction by experience gained in-situ or in comparable conditions elsewhere, or by yield projection methods, is often not possible because of the novelty of management systems and the differences between different areas of tropical forest.

SYMFOR is a software framework for ecological and management models relating to mixed tropical forests (Phillips and van Gardingen, 2001a; 2001b). It uses Permanent Sample Plot (PSP) data to simulate processes in the forest at the level of individual trees. This enables predictions to be made regarding the future of the forest under scenarios for which field data do not necessarily exist. Models within SYMFOR have been applied in Indonesia to establish the sustainability of alternative management scenarios (van Gardingen *et al*., 2001), and to test criteria and indicators of sustainability (van Gardingen and Phillips, 1998). Before this study, models such as this did not exist for Brazil, although suitable PSP data were available.

The demand for an individual-based growth and yield model for Brazil implemented within the SYMFOR framework was provided by the linkage (Phillips, 2002a) of the simulation models Eco-Gene (Degen and Scholz, 1996,1998; Degen *et al.*, 1996,1997) and SYMFOR.

The “SFC” dataset from PSP plots in various locations around the central, lowland, *Terra Firme* Amazon region was identified as being suitable for this analysis. Previous studies (Alder and Silva, 2000; Silva *et al*., 1995, Silva *et al*., 1996) have used this dataset for growth and yield studies and the development of a cohort model of tree growth and mortality.

This report describes the models developed using the SFC dataset. The approach used was a similar sequence to that used for Indonesia (Phillips *et al*., 2002a) and Guyana (Phillips *et al*., 2002b):

- Collation and assessment of data quality;
- Production of ecological species groups;
- Calibration of the diameter growth model component;
- Development and calibration the mortality and recruitment model components;
- Integration of the model components to form an ecological model in the SYMFOR framework; and
- Tune the ecological model to the assumption that primary forest doesn’t change statistically with time.

2 Model description

2.1 Introduction

The ecological model was developed to be implemented in the SYMFOR framework (Phillips and van Gardingen, 2001a, b). The model was developed to represent the processes of *diameter increment* (growth); *recruitment* (new trees appearing in the stand at the minimum diameter threshold of 5 cm); and *mortality* (death from natural causes). These processes use a representation of competition in the form of a competition index, derived explicitly for each tree in the plot. The model was designed to run with an annual timestep. A summary of the model parameters and their implementation within SYMFOR is given in Appendix A.

2.2 Competition index

Absolute competition index

The competition index, C_t, for tree t, is calculated based on the grid-square location of the trees:

$$C_t = z_1 \sum_{i=1}^{n_1} \frac{D_i}{D_t} + z_2 \sum_{j=1}^{n_2} \frac{D_j}{D_t} + z_3 \sum_{k=1}^{n_3} \frac{D_k}{D_t} \tag{1}$$

where: z_1, z_2 and z_3 are the coefficients for the relative competition importance of zones 1, 2 and 3, respectively; i, j and k are the over-topping trees in the three zones; n_1, n_2 and n_3 are the total number of over-topping trees in the three zones; and D is the tree diameter. Zone 1 is the 10 by 10 m grid-square containing tree t, and zones 2 and 3 are defined relative to zone 1 as shown in Figure 1.

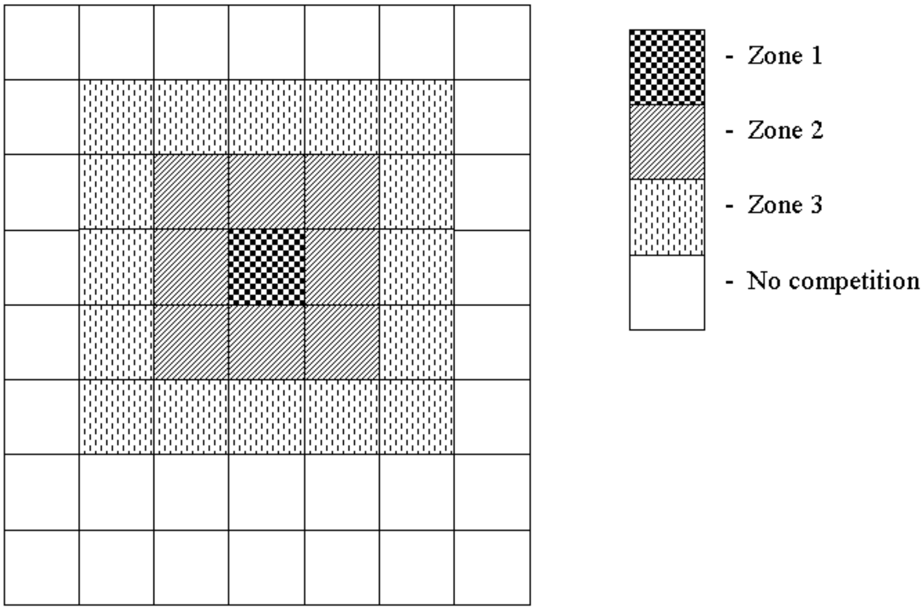


Figure 1: The different zones of competition, for a tree in zone 1. The competition indices from these different zones may be combined in different ratios, leading to competition indices that depend differently on distance from the object tree. Whatever the size of the plot, no trees located outside the three zones contribute to the competition index value.

Diameter-independent competition index

To help understand the influence of competition on growth and other dynamic tree attributes, the competition was modelled as a function of diameter, and the modelled value was then subtracted from the absolute value. This leads to a competition index symmetric about zero, and, on average, independent of tree diameter.

The diameter-independent competition index, C, is calculated by the equations:

$$\hat{C}_{\mathrm{t}} = \frac{b_0}{b_1 + D} + b_2 \qquad (2)$$

$$C = C_{\mathrm{t}} - \hat{C}_{\mathrm{t}} \qquad (3)$$

where $\hat{C}_{\mathrm{t}}$ is the predicted value of the competition index, C_{t}, calculated from the tree diameter, D, and b_0, b_1 and b_2, are parameters.

The diameter-independent competition index, C, is the competition index used in the subsequent models.

2.3 Growth model

The form of the growth model was based on the ecological concept of competition for resources between trees. Growth was defined as a function of diameter (D) and a competition index. No further ecological criteria were applied in terms of the functional form of the growth equation.

The model used for the Brazilian data is the same as that used for Guyana (Phillips *et al*., 2002b), and a simplification of the version developed for Indonesia (Phillips *et al*., 2002a).

The growth of individual trees, I, was described as the predicted diameter increment for a tree in a given year described by the equation (1):

$$I = D(a_0 + a_1 e^{-a_2 D}) + a_3 C + a_4 \qquad (4)$$

where a_0, a_1, a_2, a_3, a_4 and a_5 are model parameters and D is the DBH of the tree. The diameter-independent competition index, C (eqn. 3), was devised to describe the competition environment for each individual tree in the plot.

2.4 Recruitment model

Models of recruitment describe the appearance of new trees in the simulation at or just above the minimum diameter threshold used in the model. The SYMFOR framework does not represent small individuals (seedlings and saplings) and for this reason the ecological processes of germination, growth and mortality of seedlings cannot currently be described. An alternative approach has been implemented where the probability of a new tree becoming established is described as a function of the environment within small gridsquares within the plot. The annual probability of recruitment (F) occurring in that grid-square was modelled as a function of the mean growth rate using equation 7:

$$F = r_1\, \mathrm{e}^{(-r_2 \cdot I)} + r_3 \qquad (5)$$

where r_1, r_2 and r_3 are parameters and I is the predicted growth rate of a tree at a randomly selected location within the grid-square. The growth rate is predicted using the growth model described above (eqn. 4), for a tree with the same diameter as the minimum DBH threshold of 5 cm.

A model parameter, T_{I}, represents the time required for ingrowth as the number of years required for a tree to grow from seed to a DBH of 5 cm. It is used in the simulation when an area of ground is cleared of seedlings, for example when the soil surface is damaged during log extraction.

2.5 Natural mortality model

Natural mortality (M) is modelled as a stochastic process dependent on diameter. The probability function, or its derivative, has discontinuities at two diameters: $b_d + 5$ cm (to represent increase mortality for smaller trees), and D_{95}, the 95-percentile value of the diameter probability distribution, to represent increased mortality for trees reaching their biophysical limit. The equation used calculate natural mortality probability, M, for a tree of diameter D (cm), was:

$$M = \begin{cases} m_0 & \text{if} \quad D < b_\mathrm{d} + 5 \\ m_1 & \text{if} \quad b_\mathrm{d} + 5 \le D < D_{95} \\ m_1 + m_2(D - D_{95}) & \text{if} \quad D_{95} \le D \end{cases} \tag{6}$$

where m_0, m_1 and m_2 are parameters. The basis for this approach is discussed later (see section 3.8).

2.6 Damage mortality model

Simulated falling trees cause simulated damage and associated mortality to the surrounding trees. This is not as in other SYMFOR models (Phillips *et al.*, 2002a; 2002b), however, due to the presence of lianas in Amazonian forests. The stochastic determination of the direction of a tree's fall leads to an area in which all trees are killed by damage.

The basis of the model is that only large trees fall, but, when they do, they create a large area of damage in which all trees are killed. The size and shape of the gaps would vary, in practice, but for modelling purposes a simplification is required. The size and shape of the area of damage is based around the dimensions of the falling tree, converted to a rectangular shape. The dimensions and shape are shown in Figure 2.

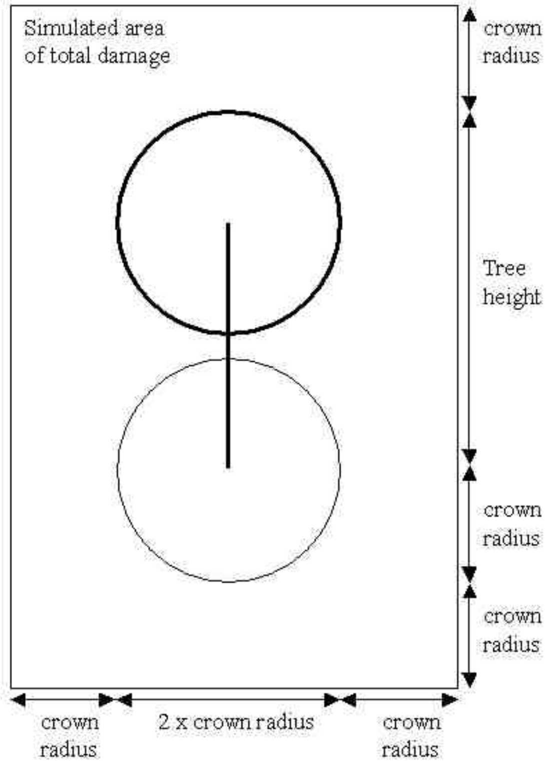


Figure 2: The shape and dimensions of the rectangular area of damage created by the fall of a large tree. The heavy lines represent the stem and crown of the tree in their positions after tree-fall, with the lighter circle representing the original position of the crown before tree fall.

The length of the area of damage is the tree height plus 3 times the crown radius. The width of the area of damage is 4 times the crown radius.

If the user elects to simulate no liana cutting before logging, damage resulting from felling operations is simulated in the same way. If liana cutting is selected, the simulation of damage created is the same as for falling trees in other SYMFOR models (Phillips *et al.*, 2002a, b) for forests where lianas are not present.

Trees above a minimum diameter, D_{m}, that die, fall with a probability of P_{f}.

2.7 Other functions

The SYMFOR framework required estimates of the dimensions of individual trees to simulate processes such as damage during harvesting. These are, total tree height H, crown-point height C_{P} and crown-radius C_{R}. The basal area, B and volume, V of individual trees is required for the management model and analysis of results from simulations. All of these attributes are derived for each tree from values of DBH, D, using auxiliary functions.

Total tree **height**, H, is calculated by an inverse linear relationship with DBH, D:

$$H = \frac{sDH_{\mathrm{m}}}{sD + H_{\mathrm{m}}} \tag{7}$$

where s has a value 200 and H_{M} represents maximum tree height with a value of 50 m.

The "**crown-point**", C_{P} (m), of a tree was defined to be the height at which the tree has maximum crown width, calculated using a simple linear relationship with tree height, H (m):

$$C_{\mathrm{P}} = f_{\mathrm{C}} H \tag{8}$$

where the parameter f_{C} has the value 0.55.

The "**crown-radius**", C_{R} (m), of a tree was defined to be the radius of the tree crown, which is assumed to be circular in cross-section, calculated using a relationship with tree diameter, D (cm):

$$C_{\mathrm{R}} = \left(\frac{10}{D} + \frac{1}{15} \right)^{-1} \tag{9}$$

The coefficients and assumptions made in equations 7, 8 and 9 are derived from anecdotal experience since data are not available for rigorous calibration.

The calculation of tree stem **basal area**, B, assumes that the stem cross-section is circular. Thus:

$$B = \frac{\pi}{4} D^2 \tag{10}$$

Stem **volume**, V, is calculated as a function of diameter, D (cm), and crown-height, C_{P}, as:

$$V = 0.5179 \left(\frac{D}{100} \right)^2 C_{\mathrm{P}} + 0.0775 \tag{11}$$

which was calibrated by Queiroz (1984). For this equation to be valid, the coefficient values used in equations 7 and 8 are required to be as stated.

3 Calibration

3.1 Introduction

Calibration is the process of calculating or estimating values for the parameters used in the model for a particular region or type of forest. Data from the “Sistema de Inventario Continuo” (SFC) database were used to calibrate the model.

The species described in the data were grouped into ecological species groups for three reasons:

1. To enable calibration of the model by reducing the number of taxa for which models were produced;
2. To enable calibration of the model by increasing the amount of data per taxa;
3. To enable the description of the forest in terms of ecological functional types.

One set of model parameters (see Section 2) were produced for each of the species groups.

3.2 Data

Details of the data are described by Phillips (2002b, c) and Silva *et al*. (1995). The physical environment is described by Silva *et al*. (1996). The plot layout and data collection information are held by EMBRAPA Amazonia Oriental.

PSP datasets have been recorded from forests in the Tapajos and Jari Florestal lands in the state of Para. They have already been used to analyse the impact of logging on species composition, growth and forest structure (Silva *et al*., 1995), and for growth and yield studies (Silva *et al*., 1996). They were also used for the definition of species groups and growth functions of the growth model CAFOGROM (Alder and Silva, 2000). The data are stored in the SFC database format, but were extracted from this format into a single-file format by the program ExtractSFC (Phillips, 2002b).

The series of plots at Tapajos km 114 comprise 60 PSP, each of 0.25 ha. All trees with DBH greater than 5 cm have species and DBH recorded. The data describing trees include information about which 10 x 10 m sub-plot they are in, but no finer spatial resolution is held. Sub-samples are made for seedlings (taller than 30 cm) and saplings (DBH greater than 2.5 cm), but growth information is not available since individuals of these sizes are not marked. 48 of the plots are located in forest that was logged-over in 1981, with 5 inventories made over the period 1981 to 1995. 12 of the plots are in unlogged forest, with 4 inventories made over the period 1983 to 1995.

The series of plots at Tapajos km 67 comprise 36 plots, each of 0.25 ha. They have the same design and inventories as the plots described above. The forest containing the plots was logged in 1979, with 7 inventories made over the period 1981 to 1997. An analysis of these plots was made in 1994 by Silva *et al.* (1994).

The series of plots in Jari comprise 40 plots, each of 1 ha. All trees with a DBH greater than 20 cm DBH were inventoried, with DBH and species information recorded. The tree data includes information about which 10 x 10 m sub-plot they are in, but no finer spatial resolution is held. 5 % sub-samples were made for trees with DBH between 5 and 20 cm. 36 plots were logged in 1985, and 4 plots have not been logged. They were all inventoried 6 times over the period 1984 to 1996.

3.3 Data cleaning and intermediate processing

The data were cleaned based on studies of the DBH observations, the taxonomic identification records, and the mortality codes. This led to some data being changed (taxonomic data and mortality code data), and some data being rejected (normally because of erroneous DBH observations).

Taxonomic data

Vernacular, or common, names (termed “vulgares” in the data), often mapped to more than one taxon. For this reason, taxonomic identifications were used during analysis. Errors were identified for situationals where two (or more) taxa had been given very similar, but different, names. Incorrect spelling was not corrected, as long as it was consistent. All names were changed to lowercase, to eliminate case-mismatches.

Family	Genus	Species	Vernacular	Incorrect field	Correction
apocynaceae	geissospermum	senceum		Species	sericeum
burseraceae	protium	sagatianum		Species	sagotianum
flacourtiaceae	casearia	javitensis		Species	favitensis
lauraceae	nectandra	micrantha		Species	myriantha
bombacaceae	quararibea			Genus	quaribea
icacinaceae	emmotium			Genus	emmotum
ochnaceae				Family	ochanaceae
ochanaceae	ouratacea			Genus	ouratea
Ni	Ni	ni	cipo	Family	cipo
" "	" "	" "		Family	ni
Ni	" "	" "		Genus	ni
Ni	Ni	" "		Species	ni
borraginaceae				Family	boraginaceae
		sp.		Species	spp
		sp		Species	spp
lauraceae	aniba	canelila		Species	canellila
lauraceae	aniba	canelilla		Species	canellila
ochanaceae	ouratea	ploygyna		Species	polygyna
sapotaceae	mycropholis			Genus	micropholis
anacardiaceae	anacarduim			Genus	anacardium
chrysobanalaceae			Family	chrysobalanaceae	
monimiaceae	siparuma			Genus	siparuna
Violacea				Family	violaceae
simarubaceae				Family	simaroubaceae
caesalpiniaceae	dialium	guianense		Species	guianensis
leguminosae	tachigalia	myrmecophyla	Species	myrmecophylla	
leguminosae	dialium	guianensis		Family	caesalpiniaceae
leguminosae	hymenaea	courbaril		Family	caesalpiniaceae
leguminosae	peltogyne	paradoxa		Family	caesalpiniaceae
leguminosae	tachigalia	spp		Family	caesalpiniaceae
leguminosae	tachigalia	myrmecophylla	Family	caesalpiniaceae	
leguminosae	dinizia	excelsa		Family	mimosaceae
leguminosae	enterolobium	schomburgkii	Family	mimosaceae	
leguminosae	inga	heterophylla	Family	mimosaceae	
leguminosae	inga	spp		Family	mimosaceae
leguminosae	parkia	gigantocarpa	Family	mimosaceae	
leguminosae	parkia	pendula		Family	mimosaceae
leguminosae	stryphnodendron	pulcherrimum	Family	mimosaceae	
leguminosae	diplotropis	purpurea		Family	fabaceae
leguminosae	dipteryx	odorata		Family	fabaceae
leguminosae	hymenolobium	excelsum		Family	fabaceae
leguminosae	ormosia	flava		Family	fabaceae

Table 1: The changes made to the family, genus and species records. Note that italics are not used since these are data replicas and are not necessarily scientific names. The changes are listed in the order in which they were made, which is important for obtaining the correct results.

The changes that were made are shown in Table 1 (in the order in which they were implemented). This process led to a reduction in the number of taxa in the dataset from 547 to 485.

Trees with recorded diameter of zero were identified as errors, and, since the mortality codes were also often zero (meaningless) for these trees, the DBH values were set to "missing". Tree that had no valid diameter records were rejected in the analysis.

The mortality code, "codcf", in the SFC data, was used to indicate more than just mortality. There were codes for living trees that described damage (or a lack of it), and additional codes for mortality and its various causes. The codes for living trees (111, 112 and 121) were merged for use in SYMFOR.

All DBH records from a survey where the mortality code indicated the tree was dead or not found were set to "missing".

Producing growth data

The survey data were used to estimate periodic growth. Typical measurement intervals were 2 years, which was usually adequate for the diameter increment to dominate the associated measurement uncertainty, and thus give meaningful data. Each interval was used to calculate a separate growth observation for each tree.

47,581 tree observations produced 142,648 growth observations. Initial analysis showed variation in growth observations between –16.0 cm and +22.3 cm. Some erroneous measurements are to be expected, and these were examined using a systematic process. It should be noted that it was not desired to remove the observations that had surprising values arising purely from measurement uncertainty. All observations are affected by measurement uncertainty and regression procedures assume the distribution of data is symmetric about the mean. Removing particular values from the distribution could therefore bias the regression. The blanket removal of negative growth observations, in particular, at this stage would lead to overestimates of growth rate by the resulting model.

The aim of the data cleaning process was to remove observations whose values were the result of specific mistakes (as opposed to "measurement error"). The following criteria were adopted for the rejection of growth observations:

- If the growth observation was missing (one or more DBH observation was missing);
- If the tree had been subject to silvicultural treatment;
- If the observed growth rate was less than –0.5 cm.yr^{-1} and it was either the first or last growth observation for a given tree;
- If either the growth observation value or the preceding growth observation value was less than –0.5 cm, and if the ratio of these two observations was between the values –0.5 and –2.0;
- If either the growth observation value or the following growth observation value was less than –0.5 cm, and if the ratio of these two observations was between the values –0.5 and –2.0.

These criteria were designed to exclude data where a single DBH observation had been recorded incorrectly, thus affecting two growth estimates. For the first or last DBH observation, this process was not possible, and a simpler approach was adopted. The effects of this rejection algorithm are given in Table 2.

Description	N
Accepted growth data	136,466
Reject first growth observation if large-negative	41
Reject last growth observation if large-negative	25
Reject growth observations if part of a (DBH) "lo hi lo", or vice versa, sequence (one DBH observation wrong)	929
Reject ringed or otherwise treated trees	5,187

Table 2: The numbers of growth observations rejected, with the reason for rejection.

Producing Ingrowth data

Ingrowth (recruitment) past a diameter threshold of 5 cm was identified by a missing diameter in one survey followed by a non-missing diameter in the next survey. Only data from the Tapajos regions were used in this analysis, since these were the only datasets for which trees with a diameter of 5 cm were recorded.

Two criteria were applied to exclude data:

- If the potential recruit tree had a previous non-missing observation, suggesting that the candidate recruitment observation was an artefact of missing data;
- If the growth rate in the interval following recruitment was observed to be at least 5 times slower than the (minimum) growth rate observed during recruitment, assuming the recruit had a diameter of 5 cm prior to initial observation, suggesting that the candidate recruit was an artefact of missing data.

38,939 tree observations from the Tapajos regions, collected over a period of 16 years, led to 14,002 observations of recruitment. 28 observations were rejected because of the above two criteria. 13,974 observations were accepted after the criteria were applied.

Producing mortality data

Tree mortality was identified by trees with a non-missing DBH observation in one survey, followed by a missing DBH observation. This process identified 14,398 observations of mortality from 47,581 tree observations. Three of these observations had a mortality code representing “living tree”, and were assumed to be natural mortality. The breakdown of mortality by mortality code is given in Table 3.

Number of trees	Description – cause of death
10598	Natural causes
818	Logging operations (not logging)
716	Silvicultural treatment
350	Logged
14	Reason for death unknown
1902	Tree not found

Table 3: Breakdown of mortality data by mortality code.

It was assumed that the trees that were not found had died from natural causes, bringing the total number of observations of natural mortality to 12,500.

3.4 Competition model

The values used for the parameters z_1, z_2 and z_3, in equation (1), calculating the absolute competition index for each tree, were 9, 4 and 1, respectively. These values were selected arbitrarily to represent the decrease in competition with increasing distance.

The parameters in equation (2), the model of absolute competition index as a function of tree diameter, were evaluated by regression. The values of these parameters are given in Table 4.

Parameter	Value
b_0	6004
b_1	-1.580
b_2	-130.8

Table 4: The values of the parameters in the equation modelling absolute competition index as a function of diameter.

3.5 Species groups

In order to group the species, it was necessary to assess the population density for each taxon recorded in the data set. Table 5 shows that 85 taxa were represented by more than 100 trees. This number progressively increased as the minimum number of trees decreased.

Minimum number of trees	Number of taxa	Number of trees	% of sample
100	85	38848	81.6
50	143	43064	90.5
20	235	46107	96.9
1	485	47581	100.0

Table 5: Minimum number of trees in a taxon for the permanent sample plots.

The process for grouping species involved three stages: (1) a clustering analysis to make the groups using the most populous species; (2) discriminant analysis to add the less populous species to the existing groups; and (3) a subjective stage where species with little or no data were assigned to the groups.

Clustering analysis of populous species

For each species, a set of variables was produced:

1. Mean growth rate;
2. Average growth rate at low competition;
3. Average growth rate at medium competition;
4. Average growth rate at high competition;
5. Average growth rate of new recruits (DBH of 5-6 cm);
6. D_{95}, the 95-percentile point in the DBH frequency distribution (as an index of mortality behaviour).

Low and high competition levels were specified using the diameter-independent competition index C (eqn. 6). Values of C above 40.0 were classed as being *high competition*, and values below –40.0 were classed as being *low competition*, with *medium competition* being defined between these values ($-40 < C \leq 40$).

The grouping process used taxa with at least 50 trees, giving 143 species (Table 5). 40 species had missing data for at least one of the 6 characteristic variables, and could not be included in the clustering analysis.

Thus only 103 species were used at this stage. Data were evaluated for these species, and normalised so that the range of values of each variable was from 0.0 to 1.0. A clustering procedure was then used to group the species according to the normalised values of variables 2 – 6 (above).

The clustering process requires the user to decide how many groups there should be in advance. Ten groups were used to define the initial grouping to be consistent with previous studies (Phillips *et al*., 2001). The groups and their normalised variables were then examined. The production of 10 groups was able to separate species with extreme ecological characteristics, and discriminated between the species at opposite ends of the range of characteristic values. The number of groups was retained for the next step of the species grouping process.

An alternative clustering process was applied using the values of variables 1 and 6 (above) only. This process led to a similar result as for variables 2 – 6, and was adopted since it was simpler.

The discriminant analysis: for less populous species

Remaining ungrouped species were added to the existing groups using discriminant analysis. The data from the existing groups were used as training data to initialise this process. The characteristics used to define the groups and the species were the 95 percentile point on the cumulative DBH distribution (D^{95}) and the mean growth rate (as for the clustering process).

Group	Name (reference)	D^{95}	Growth rate	Dominant members	N trees	N species
1	Medium-growing mid-canopy	41.8	0.21	Sapotaceae, Lauraceae*, Guatteria poeppigiana*	8978	87
2	Slow growing understorey	15.9	0.09	*Rinorea flavescens, Duguetia echinophora, Talisia longifolia*	5964	48
3	Medium growing mid-canopy	57.2	0.29	*Geissospermum sericeum, Carapa guianensis, Pouteria spp.*	4851	76
4	Slow growing lower canopy	27.7	0.18	*Protium apiculatum, Rinorea guianensis, Neea spp*	11016	101
5	Medium growing upper canopy	72.5	0.26	*Couratari oblongifolia, Minquartia guianensis,*	2732	34
6	Fast growing upper canopy	76	0.54	*Sclerolobium chrysophyllum, Trattinickia rhoifolia, Didymopanax morototoni,*	1356	29
7	Fast growing pioneers	35.8	0.54	*Inga spp., Sloanea froesii, Bixa arborea, Jacaranda copaia,*	8606	49
8	Emergents - climax	104.0	0.37	*Manilkara huberi, Goupia glabra, Hymenaea courbaril, Dipteryx odorata*	1409	36
9	Very fast growing pioneers	38.7	1.26	*Cecropia sciadophylla, Jacaratia leucoma, Jacaratia spinosa*	2007	9
10	Very fast growing upper canopy	78.2	0.94	*Tachigalia myrmecophylla, Sclerolobium tinctorium*	662	16

Table 6: Summary of the characteristics of the species groups. It should be noted that the data that were used to produce this table include data from recently logged-over forest.

Subjective grouping analysis

The species with less observations are generally less important to forest structure and dynamics, less important from a commercial point of view, and were more likely to be assigned to an inappropriate group due to being misrepresented by the variable values calculated from the dataset. In order to correct this, it is necessary to apply expert knowledge of the species formed from a general accumulation of experience over years of forestry. This was not done since suitable expertise was not available at the time of model development, but can be carried out in the future for less common species without losing model validity.

Summary

The species grouping resulting from the above process groups 485 species into 10 groups, whose average characteristics are given in Table 6.

3.6 Growth model

The purpose of the growth model was to explain some of the variation of growth rate within each species group. The model was calibrated separately for each species group using multivariate regression. The amount of variation in the dataset that was explained by the model was described by the R^2 statistic (Table 7). When combined, the species grouping and the growth model together explained 33.7 % of the observed variation in the dataset.

The residual variation in growth rates (66.3 %) was analysed to establish if any casual effects could be determined from the data. An estimate of the measurement uncertainty on any growth observation was made (using data from Indonesia, see Phillips and van Gardingen, 2002), and found to be 0.3 cm. The variation that could be explained by the measurement uncertainty was highly correlated to the mean growth rate of a species group: all of the remaining variation in growth rates of trees in slow growing species groups could be explained by measurement uncertainty, but not for species groups typified by high growth rates.

Lianas can be a dominant feature of Amazonian forests, with up to several on a single tree. They can grow to significant diameters. They initially grow up one tree, but then may spread through the canopy to affect other trees. The extent and effect of their interaction with tree crowns is not well known. It was hypothesized that they may have some effect on tree growth, as well as the process of gap formation and regeneration. The data contained a simple index of infestation by lianas for each tree. A simple model of the effect of lianas on growth rate was constructed, and could explain 1 % of the variation in growth rates. There were additional difficulties in developing a model to represent the effects of lianas, including:

- Half of the growth observations in the data set did not have usable records of liana presence;
- The data describing lianas were assigned one of three codes, and it was unclear exactly what these meant;
- Serial records for individual trees showed the liana index changing without apparent consistency;
- Within the data set there were no records that could be used to infer liana dynamics, and thus it could be applied only as a random factor contributing to the simulated variation in growth rates.

With these difficulties, and the relatively small effect on growth rates, it was decided not to include a representation of liana infestation for the purposes of growth modelling.

It appeared that the residuals from the growth model were randomly distributed and it was assumed that this variation resulted from effects including measurement error, the effect of lianas, genetic effect and site specific effects and events such as pests, diseases and weather.

Group	a_0	a_1	a_2	a_3	a_4	R^2 (%)
1	0.0025	0.0077	0.0452	-0.0005	0.1174	4.9
2	0.0054	-0.0111	0.0504	-0.0002	0.1136	5.7
3	0.7447	-0.7389	-0.0001	-0.0007	0.1953	3.3
4	0.0041	0.0211	0.1590	-0.0005	0.1030	5.5
5	0.0014	0.0221	0.0291	-0.0004	0.0023	8.0
6	-0.0506	0.0811	0.0046	-0.0009	0.1523	14.4
7	2.2632	-2.2464	-0.0001	-0.0016	0.3680	13.0
8	-0.0029	-0.1747	0.1634	-0.0008	0.6001	9.2
9	0.0093	0.3329	0.0776	-0.0050	-0.4276	18.3
10	0.0065	0.0805	0.0287	-0.0011	-0.1354	15.9

Table 7: Parameters for the growth model (eqn. 4) and the associated goodness of fit R^2 (%).

3.7 Recruitment model

The probability of ingrowth was estimated at the scale of individual 10 by 10 m gridsquares as a function of diameter independent competition index *C*. The competition index was calculated for a “virtual” tree with *D* of 5 cm at the centre of each gridsquare. This was used to calculate the predicted growth rate associated with a particular species group and grid-square, using the model calibrated for growth (eqn. 4, section 3.6). The growth rate for all gridsquares was binned into a histogram for each species group.

Data from the permanent sample plots from Tapajos (for which data about trees with DBH above 5 cm were recorded) were analysed for each species group to record the number of observations of ingrowth by grid-square, and hence the number of observations of ingrowth for each growth-rate bin. The probability of ingrowth, *F*, was then calculated for each growth rate bin for each species group:

$$F = \frac{N_{\mathrm{F}}}{N_G} \tag{12}$$

where N_{F} is the number of observations of ingrowth and N_{G} is the number of gridsquares. The probability of recruitment was modelled for each species group by regression as a function of predicted growth rate, *I*, using eqn. 5. The data points used for regression (one for each combination of species group and growth rate bin) were weighted by the total number of grid-squares in each bin to give the correct arithmetic mean.

The resulting regressions produced value for R^2 ranging from 0 %, for species exhibiting no ingrowth probability response to growth rate, to 86 %, for the pioneer-type species.

No data were available to estimate T_{I}, the ingrowth time parameter. The predicted diameter increments, *I*, of trees with DBH of 5 cm and a diameter-independent competition index of -250 (very low competition in unlogged forests) were used in an estimation of the “ingrowthtime” parameter T_{I}:

$$T_{\mathrm{I}} = \frac{5}{I} \tag{13}$$

This estimate is subject to a significant uncertainty since it is not based on data. Alternative methods of calculation, such as the mode or geometric mean would be equally applicable.

Group	r_1	r_2	r_3	R^2	T_I
1	0.029	-3.05	0.0065	68	18
2	0.045	2.4E-11	0.022	0	33
3	1.22	-0.036	-1.214	48	12
4	0.30	-0.82	-0.238	75	17
5	0.70	-0.02	-0.692	24	25
6	0.0079	-2.81	-7.5E-3	51	9
7	0.148	-2.29	-0.304	74	6
8	2.6E-04	-6.91	0.0045	29	13
9	0.118	-0.79	-0.275	86	3
10	0.20	-0.013	-0.197	12	10

Table 8: Parameter values for the recruitment model (eqn. 7, section 2.4).

3.8 Mortality model

The mortality probability was modelled as a function of diameter. The calibration was derived directly from the observations of mortality in the data.

To obtain the mortality probability as a function of diameter, it was necessary to split the data into diameter classes, and to count the number of mortality observations and the number of live-tree observations for each class. This was performed for each species group. Diameter class widths were allowed to vary, within some limitations. A class boundary was forced at a diameter of 20 cm, since the Jari data had a minimum DBH of 20 cm. The smallest diameter class had a lower bound at 5 cm. The upper bound of the largest diameter class was 1.5 times the size of the largest mortality observation, for each species group. The mean value of tree diameter, *D*, was used to describe each bin rather than the mid-point as the diameter distribution was not linear.

Mortality observations were made over intervals not equal to 1 year. To convert these observations into annual mortality observations, an annual-equivalent variable was used:

$$A = \frac{1}{Y} \tag{14}$$

where *Y* is the number of years in the interval over which the mortality observation was observed. For each plot, interval and species group, *A* was summed and expressed as a proportion of the number of live trees, obtaining a probability.

For all the groups, the probability of mortality was approximately constant over all diameters, with the exception of the first diameter class. The mortality was thus modelled with three parameters: the width of the first diameter class, the probability of mortality in that diameter class, and the probability of mortality in all other diameter classes.

Model tuning was performed using the assumption of dynamic equilibrium in primary forest (see section 3.9). For this model, as with other models (Phillips and van Gardingen, 2002; Phillips *et al*., 2002), the performance relating to large trees was found to be poor. This is unsurprising, since little data were available to describe the behaviour of mortality of large trees. Consequently, mortality probability is increased in the model for trees with DBH above a specified threshold (the D_{95} variable in eqn. 6) by a percentage per cm DBH, m_2. The resulting values of the parameters are shown in Table 9.

Group	m_0	m_1	m_2	D_{95}	b_d
1	2.3	1.8	0.0	41	7.5
2	2.9	2.9	0.5	15	3.75
3	1.85	1.1	0.2	57	15
4	2.9	2.4	0.5	27	5
5	1.9	0.5	0.25	72	7.5
6	4.0	2.1	0.5	76	7.5
7	6.7	5.0	0.0	35	7.5
8	3.0	0.85	0.0	104	15
9	8.9	4.5	0.5	38	5
10	4.3	4.0	0.5	78	15

Table 9: Parameters for the mortality model, equation 6, section 2.5.

3.9 Damage mortality model

Model simulations with un-calibrated, tree falls showed that any gap formation due to tree fall led to the ingrowth of more pioneers (trees of species group 7 and 9) than were observed in the PSP, even in the PSP with the most pioneers. This led to the conclusion that the PSP were perhaps (deliberately, or by chance) located in forest without significant gaps. This is possible, because the plots are small, and field experience shows the space between the gaps to be relatively large.

An analysis of the data from the Dendrogene project's Intensively Studied Plots (ISP) was undertaken in an attempt to establish the frequency and size of gaps. The ISP block 6 comprises a square 100 ha of forest, in which the positions of all trees with a DBH greater than 20 cm are recorded. One enumeration of this plot has been made.

The area was divided into 10 m by 10 m grid-squares, and those with zero tree records were selected. A clustering algorithm was used to group the grid-squares together using the position variables (x and y). For each cluster, the area was calculated using the minimum and maximum x and y values. This gave a maximum value for the area, assuming the area was rectangular. The proportion of the area of each cluster made up by zero-tree grid-squares was then calculated. Some clusters had a significantly higher proportion of gap than others. Those with over 50 % were selected.

This process selected 10 clusters, with an average area of 0.4 ha, that were more than 50 % made up of empty grid-squares that did not contain trees above 20 cm DBH. The conclusion is that 5 % of the area is in 0.4 ha areas in which 50 % or more of the 10 m by 10 m grid-squares have no trees above 20 cm. Gaps have no trees above 20 cm, however, and the area of the cluster was known to be a maximum. Thus it is reasonable to refine the previous conclusion to the statement that, "about 2 % of the area is in 0.2 ha areas with 100 % of the grid-squares having no trees with diameter above 20 cm".

The possibility of observing this from a random distribution of trees was tested. Artificial positions of the 16,111 trees observed in the ISP were generated according to a uniform random distribution throughout the ISP, and the analysis was repeated. This process was repeated 10 times. The mean "proportion of cluster area made up by empty grid-squares for the cluster with the largest value" was 0.52 ± 0.01, and the mean "number of clusters with that proportion greater than 0.5" was 1.0 ± 0.2. Since 10 such clusters were observed in the ISP, it was concluded that the trees are not distributed among the grid-squares randomly, and that real forest gaps were being identified by the analysis.

In order to model gap creation processes, the number, size and persistence of gaps was required. No data were available to evaluate gap persistence, however, so various hypotheses were tested. If a gap exists for 5 years, that means over a period of 5 years 10 gaps of 0.2 ha appear in 100 ha, which corresponds to 2 gaps per year, or 1 per 50 years per ha. Similarly, if a gap exists for 10 years that means 1 per 100 years per ha, and if a gap exists for 2 years, that means 1 per 20 years per ha. Since these gaps are defined as being the absence of trees of diameter 20 cm or more, the time for a gap to disappear is the time until a tree within the gap reaches 20 cm. For the pioneer species this was calculated as a minimum of 11 years. This suggests 1 gap per 110 years per ha, which was used to tune the model.

The area of damage, in this model, for a tree of diameter 80 cm, is 0.13 ha (as calculated by the method described in section 2.6). This is taken as a minimum diameter for a tree to fall and create a gap, leading to an average gap size of around 0.2 ha every 110 years.

It should be noted that the average area of damage observed in the ISP plot, and simulated by this model, is 0.2 ha. To avoid the edge effects dominating simulations, the minimum recommended plot size is 1 ha for simulations using this model.

3.10 Volume model

A lack of available tree stem volume data precipitated a resort to other sources. The TREMA (Thompson, 1997) software, being developed for operational forest management by the Dendrogene project (http://www.cpatu.embrapa.br/dendro/), contains two algorithms for calculation of stem volume:

$$V = \begin{cases} -0.0994 + 0.00091941 D^2 & \text{if } D < 45 \\ e^{-7.6281 + 2.1809 \cdot \ln(D)} & \text{if } D \geq 45 \end{cases} \tag{15}$$

which was developed by Silva and Araújo (1984) and Silva *et al.* (1984), and

$$V = 0.5179 \left(\frac{D}{100} \right)^2 H + 0.0775 \qquad \text{(16 and 11)}$$

which was developed by Queiroz *et al.* (1984).

These two algorithms were compared for the full range of stem diameters expected in SYMFOR simulations: 5 cm to 200 cm DBH. The height, H, in Queiroz' equation, was calculated using the SYMFOR equation for crownheight, C_p (eqn. 8), which in turn uses the equation for tree height (eqn. 7).

where D is the tree diameter (m), the parameter s has the value 200 and the parameter H_m represents the maximum height of a tree, typically 45m. Figure 3 shows the calculated volume as a function of the stem DBH for the models of Silva and Queiroz, using three different values for the height parameter, H_m: 40, 45 and 50 m. Clearly the two models agree best, and almost exactly, for the value H_m = 50 m.

The difference between the models of Silva and Queiroz, for the value H_m = 50 m, is shown in Figure 4. The discontinuity at a DBH of 45 cm in the model of Silva is clear, and despite the shape, the magnitude of the difference between the models is very small (always less than 0.5 m^3).

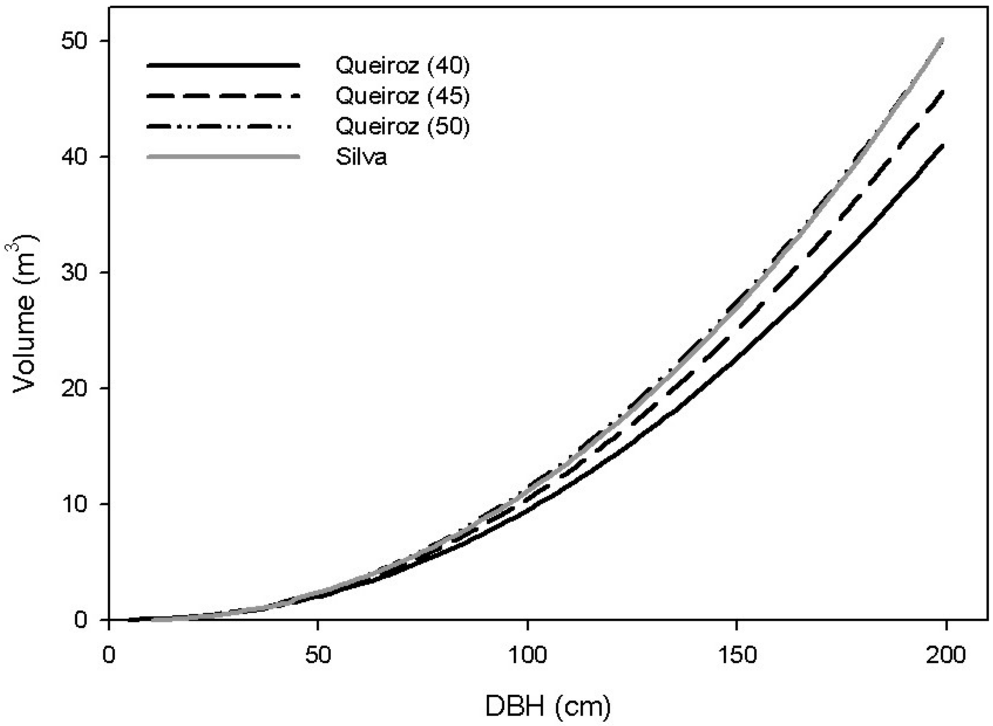


Figure 3: The relationships between stem diameter and volume, as calculated by Silva and Queiroz, the latter using three different values for the height parameter, H_m: 40, 45 and 50 m.

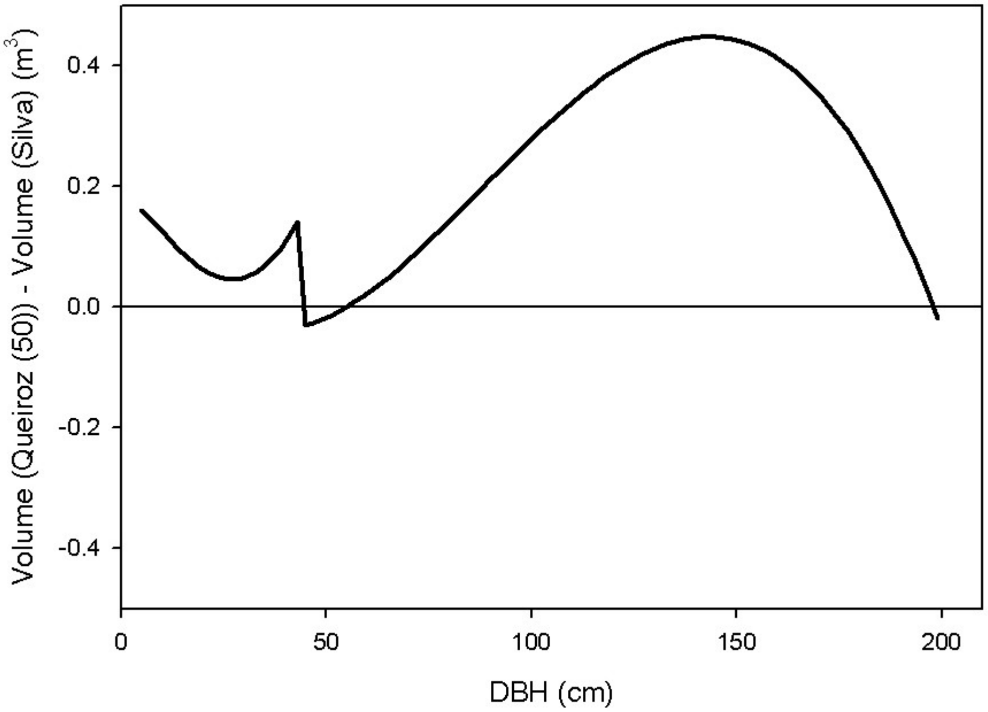


Figure 4: The difference between the models of Silva and Queiroz, for the value H_m = 50 m, as a function of stem diameter.

The discontinuity at a DBH of 45 cm in the model of Silva is not ideal, and the model of Queiroz was chosen in preference for use in the SYMFOR model, with parameter values for crown-point and tree height as given in section 2.7.

3.11 Tuning the model

The "Brazil-Tapajos" model described above was implemented in SYMFOR framework. Initial analysis of the model demonstrated that some of the parameters needed to be altered in order to achieve adequate performance. Selected model parameters were modified (tuned) to improve performance in order to demonstrate a dynamic equilibrium in primary forest.

The parameters that were varied as a result of this process were those considered to have the largest uncertainty resulting from the process of calibration:

- Mortality probability slope parameter for trees with large DBH values (m_2);
- Mortality probability for small trees (m_0);
- Mortality probability constant component (m_1);
- The recruitment probability (r_1 and r_3);
- The recruitment as a function of growth rate, for pioneer species (r_1).

The latter change was required since, for the same value of competition index, the pioneer species groups showed strong recruitment in logged-over forest compared to unlogged forest. This suggests that the competition index does not perfectly represent the conditions that affect trees growth following logging. It is suggested that the lack of absolute spatial resolution, necessitated by the lack of absolute tree-position data, may be the cause of this, since the effect has not been seen in other models (Phillips and van Gardingen, 2002; Phillips *et al*., 2002). Another possible influence is the lack of data from unlogged forest. Of the PSP, only 7 ha represent unlogged forest dynamics out of a total of 64 ha of data. The model was thus primarily calibrated from logged-over forest, but then tuned to the unlogged forest scenario.

Changes to these parameters were made incrementally over many repeated simulations. The parameters of the model demonstrated significant interactions often resulting from the underlying ecological nature of the model.

The model parameters shown in Tables 4 and 5 show the values after these modifications.

3.12 Evaluation and validation

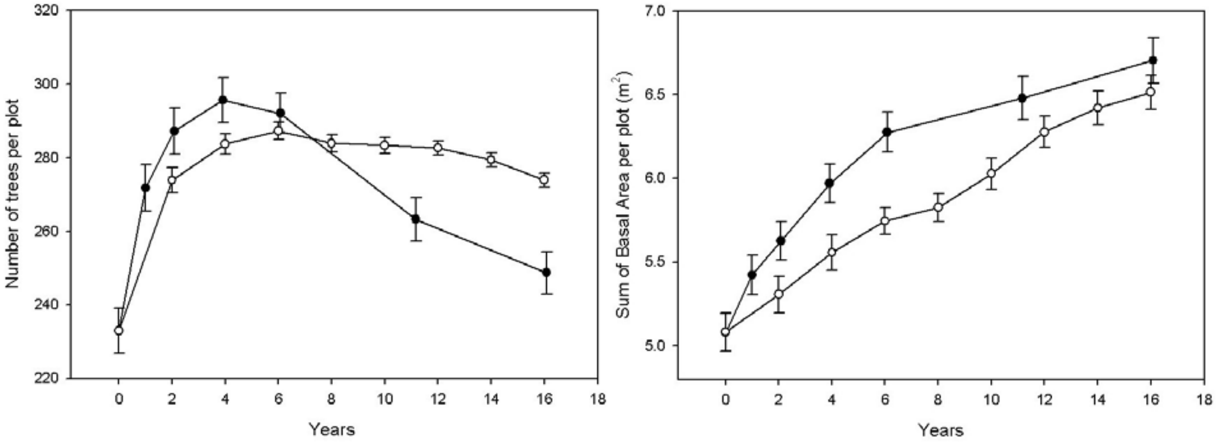
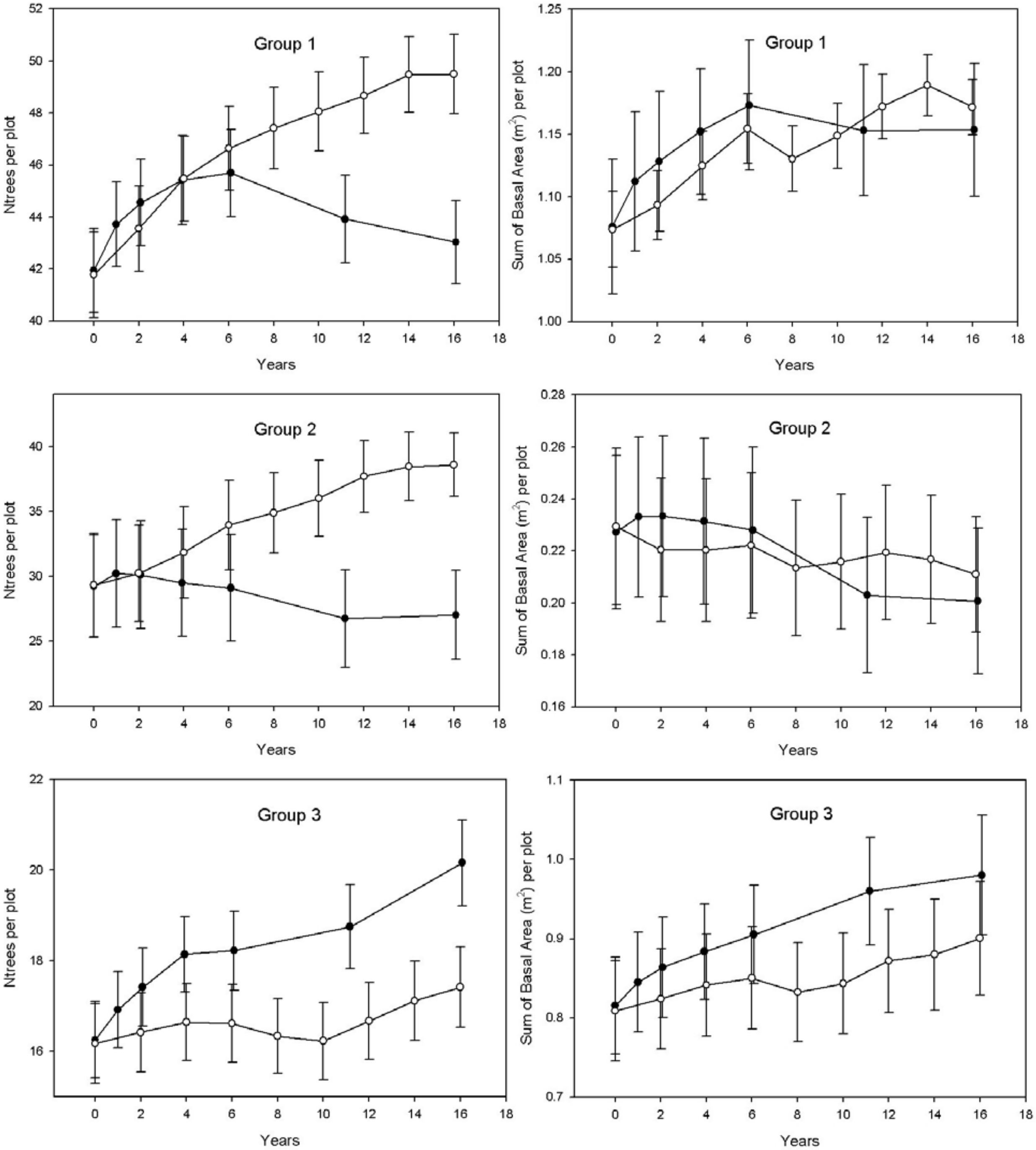
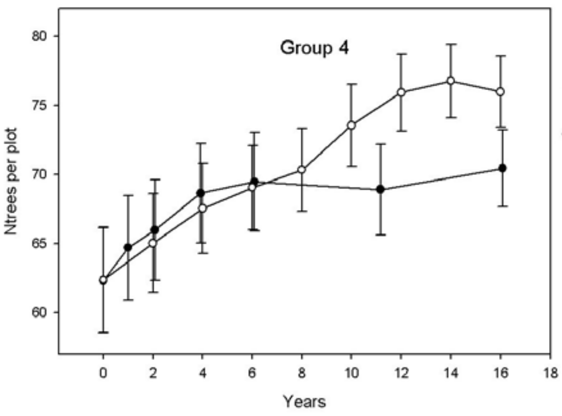
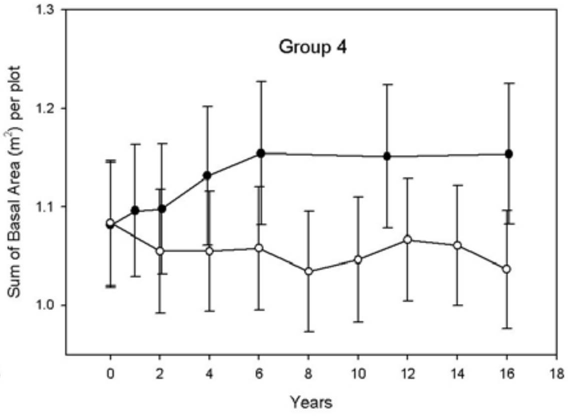


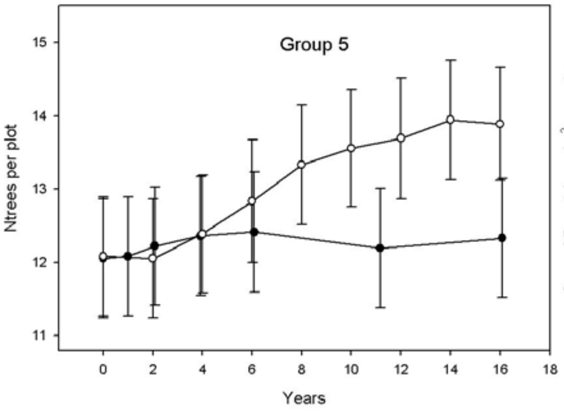
Figure 5: The variation in the mean number of trees, and the mean sum of basal area, per plot for the Tapajos km 67 plots (starting two years after logging) for the observed data (•) and the simulation model (o). The error bars are the standard error on the mean over all plots, and therefore arise from the variation between plots.

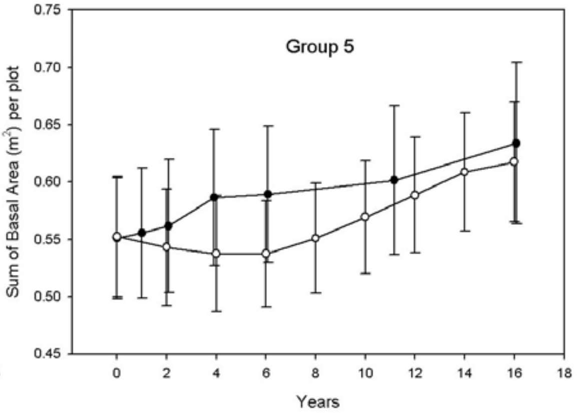
The performance of the model was evaluated using a long time-sequence of real data. The longest time sequence available is 16 years, for the plots from Tapajos km 67. These plots are described by Silva *et al.* (1995). They were logged in 1979, and the plots were set up in 1981. The most recent enumeration of the plots was made in 1997. There are 36 plots, each 0.25 ha. The data from 1981 were used to initialise simulations that were run for 16 years, after which the simulated forest data were output to be compared with the real data from the most recent enumeration of the plots. Results are shown for all species groups in Figure 5 and for each species group in turn in Figure 6.

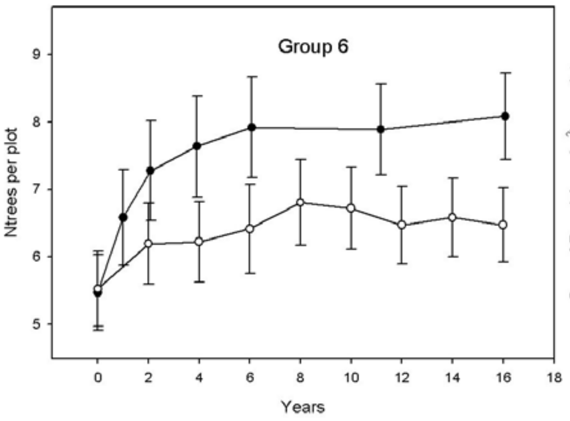


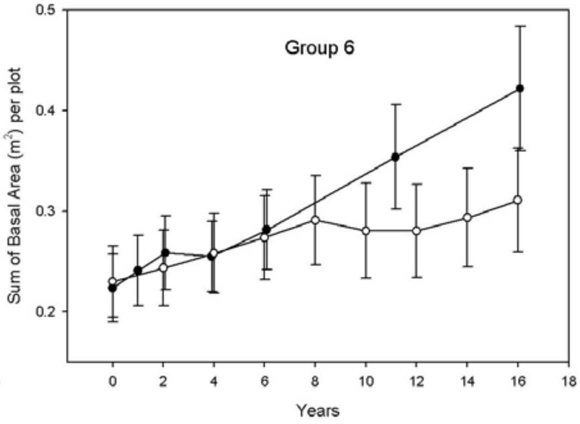

Group 4
Ntrees per plot
80
75
70
65
60
0
2
4
6
8
10
12
14
16
18
Years

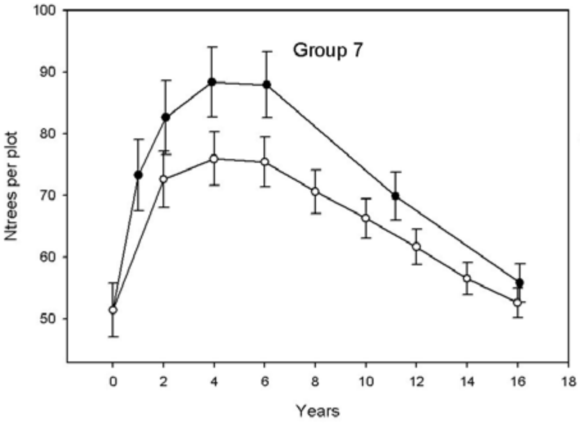

Group 4
Sum of Basal Area (m^2) per plot
1.3
1.2
1.1
1.0
0
2
4
6
8
10
12
14
16
18
Years

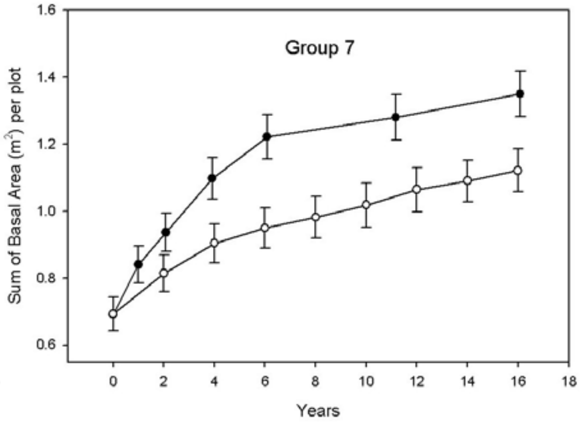

Group 5
Ntrees per plot
15
14
13
12
11
0
2
4
6
8
10
12
14
16
18
Years


Group 5
Sum of Basal Area (m^2) per plot
0.75
0.70
0.65
0.60
0.55
0.50
0.45
0
2
4
6
8
10
12
14
16
18
Years


Group 6
Ntrees per plot
9
8
7
6
5
0
2
4
6
8
10
12
14
16
18
Years


Group 6
Sum of Basal Area (m^2) per plot
0.5
0.4
0.3
0.2
0
2
4
6
8
10
12
14
16
18
Years


Group 7
Ntrees per plot
100
90
80
70
60
50
0
2
4
6
8
10
12
14
16
18
Years


Group 7
Sum of Basal Area (m^2) per plot
1.6
1.4
1.2
1.0
0.8
0.6
0
2
4
6
8
10
12
14
16
18
Years

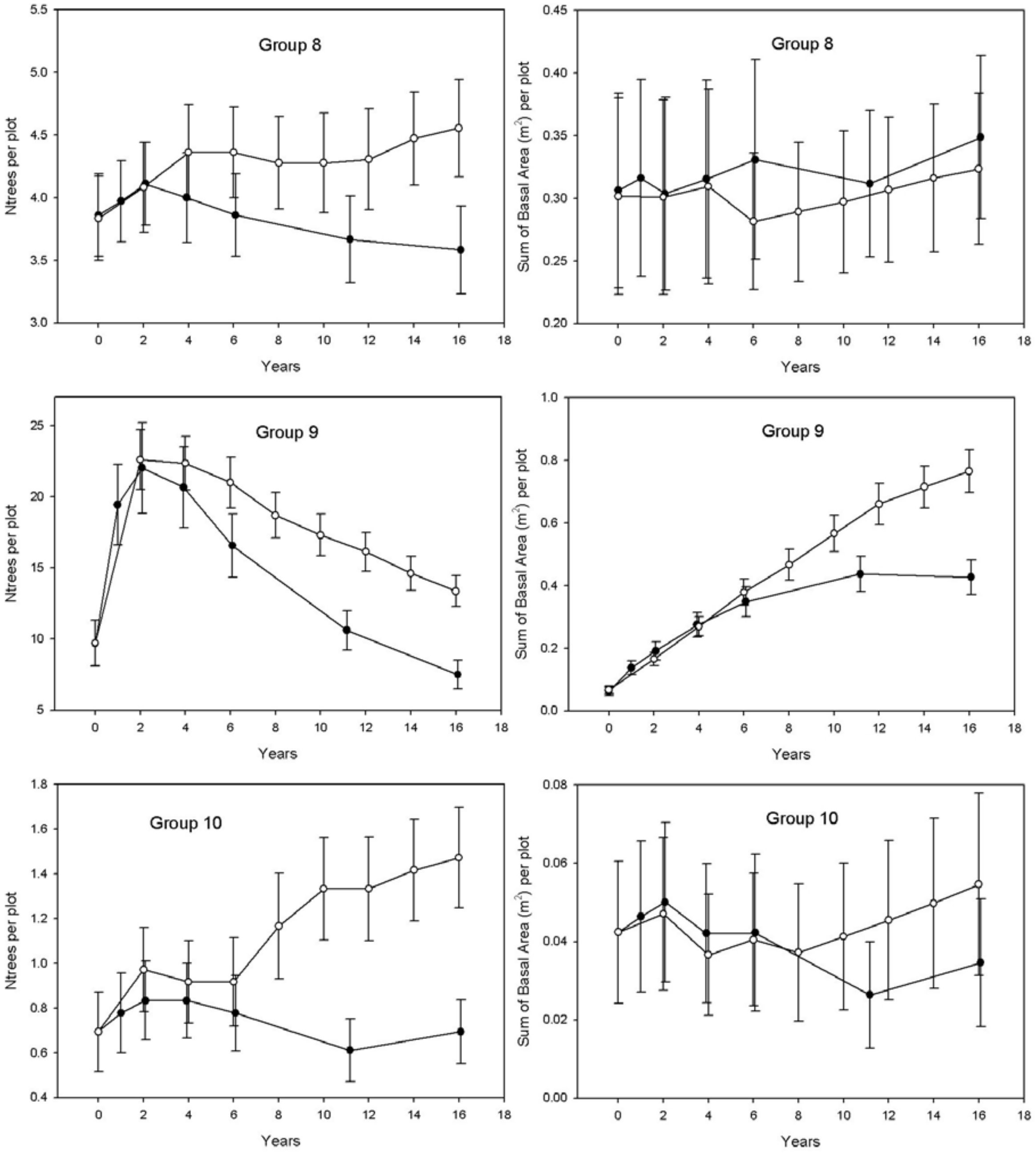


Figure 6: The variation in the mean number of trees, and the mean sum of basal area, per plot for the Tapajos km 67 plots (starting two years after logging) for the observed data (•) and the simulation model (o), shown for each species group separately. The error bars are the standard error on the mean over all plots, and therefore arise from the variation between plots.

The model was tuned to the situation of unlogged forest, but validated against a sequence of data recorded immediately following heavy logging. This is a serious test of the performance of the model; the conditions in the logged-over forest vary considerably from the unlogged forest state, and these conditions are represented in the model by a simple competition index.

The time period of 16 years was sufficient to cover much of the complex forest dynamics following a heavy logging event. The general trends of the model reflected those of the data well (see Figure 5). For many species groups, the simulated behaviour, in terms of the number of trees and sum of basal area, agreed well with the real data (see Figure 6).

A notable exception to this was the number of trees after year 8 for many species. Following logging, the reduced competition in many areas led to increased recruitment of most species groups. Perhaps, in the real forest, there was over-crowding of the gaps by small trees, leading to increased mortality of the small trees in the following years. The model does not include dependence of competition on the probability of mortality (although this relationship was tested during model development), and increased mortality in the small trees was not observed. The growth of these small trees will be inhibited by the competition so that, even while the number of trees was over-estimated, the basal area is well modelled at the end of the validation period.

Other limitations in the ability of the model in this particular case should be born in mind when considering the results. (1) During and after simulated logging, the model is used to simulate damage to areas, including clearance of seedlings and saplings, which will take some time to regenerate. The simulation described here was initialised with data from forest that was already logged over, and thus the removal of seedlings and saplings was not simulated. In practice the gaps will exist for longer than simulated, because SYMFOR assumes that, in the absence of simulated cleared gaps, seedlings exist. (2) The plots used in this study were 0.25 ha, which is the same size as the field of influence used in competition modelling (see Figure 1). Since plot-wrapping is used to avoid edge effects, the competition is likely to be almost homogeneous across each plot, and a lot of the spatial elements of the competition cannot be used.

There may be concern over the use of the same data for model development as for model validation. The situation is not as simple as this, however:

- Tuning has a strong effect on the balance between growth, recruitment and mortality for the different species groups in the model. The model was tuned to the situation of unlogged forest, but validated for logged over forest. The similarity between the data and simulation results show that the basic model form, balanced by tuning, represents the forest dynamics well.
- The sub-models of growth, recruitment, mortality and logging were developed separately, tuned to work together for unlogged forest, and validated together on logged-over forest.
- The data used for validation were a sub-set of the data used for model calibration, from one particular region. The model may be expected to represent the average behaviour of all regions used for calibration, and not to represent a single region, although this was the case for validation and would normally be the case in a later analysis.

Given these considerations, the validation is expected to be a rigorous test of the model.

4 Discussion

The characteristics of the Brazil-Tapajos model are discussed in terms of the characteristics of the species groups derived from the data and a series of figures demonstrating the performance of the resulting model in primary forest and after simulated logging.

4.1 Species analysis.

The derivation of species groups (see section 3.5) used taxa based on scientific names. Some of the more significant taxa are discussed here. It should be noted that 57 of the 64 ha of forest represented by the data are from logged-over forest. The numbers of trees discussed below represent all trees observed over the whole period of data collection.

- No single taxon dominates the forest. Of 47,581 tree observations, the most populous taxon is the genus *Inga* (Mimosaceae), with 3,541 or 7.4 %. This could be partly because the trees have not been identified to species level. The most populous single species is *Protium apiculatum* (Burseraceae), with 2,134 observations, 4.5 %.
- Three of the 10 most populous taxa: *Inga sp.* (Mimosaceae), *Sloanea froesii* (Elaeocarpaceae) and *Bixa arborea* (Bixaceae), are found in species group 7. These were observed to recruit very strongly following logging, which explains their predominance in this dataset. Also in the 10 most populous taxa is *Cecropia sciadophylla* (Moraceae), which is in the pioneer group, 9. These are indicators that the forest was heavily disturbed.
- The largest trees belong to the climax species group, 8. These include *Manilkara huberi* (Sapotaceae), *Hymenaea courbaril* (Caesalpiniaceae), *Dipteryx odorata* (Fabaceae), *Goupia glabra* (Celastraceae). The latter of these was unexpected and may be revised when the grouping is reviewed by local experts. It has been mentioned that this species is found mostly in the Jari dataset, for which data describing small trees were not used, possibly leading to misclassification by the statistical routines.
- Group 4, 1 and 2 were the first, second and third most populous species group representing 55 % of the total population. These groups represent the lower canopy, mid-canopy and understorey, respectively, and have very low growth rates, even following logging. The groups include the populous taxa *Protium apiculatum* (Burseraceae), *Rinorea flavescens* (Violaceae), *Rinorea guianensis* (Violaceae), *Guatteria poeppigiana* (Annonaceae) and the unidentified trees in the Sapotaceae and Lauraceae families.
- The pioneer species are found in groups 7 and 9. They grow at an average of over 1 cm per year in disturbed forest conditions, responding dramatically to changes in competition level.

The ecological characteristics of the species and resulting species groups could be interpreted as representing the ecological characteristics of various ecological functional types, including pioneer, canopy emergent, sub-canopy, light demanding, and shade tolerant species.

4.2 Simulations of primary (unlogged) and logged forest.

The Brazil-Tapajos model was tested in the SYMFOR framework. These applications are illustrated here.

The predicted growth rates for each species group are shown as Figure 7. Groups 1-5 and 8 are slow growing species. The highest growth rates were predicted for group 9 which is dominated by the genus *Cecropia*. Maximum growth rates for most species were observed at stem diameters of less than 40 cm.

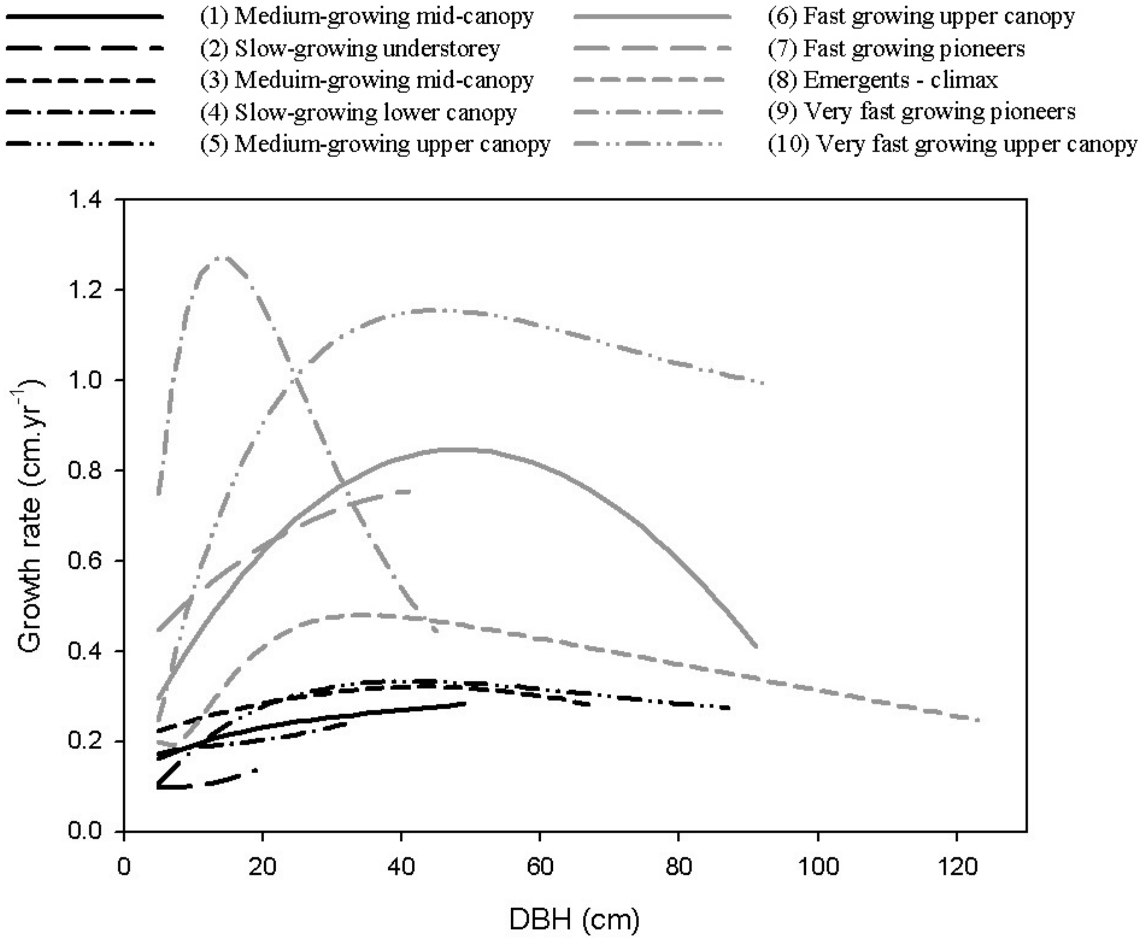


Figure 7: Predicted growth rates as a function of DBH. These rates were calculated for average diameter-independent competition (*C* = 0). The values are shown for the range from the minimum diameter of 5 cm up to the 95 percentile of the diameter frequency distribution plus 20 % for each species group.

The performance of the model in simulating forest behaviour over long timescales was illustrated by examining changes in the basal area and number of stems in each species group over time for unlogged primary forest. Figure 8 shows one example simulation reporting the total basal area of the plot (including all species groups). The total basal area shows some inter-annual variability, but the trend is that total basal area remains relatively constant.

Ecological models interact with management models in the SYMFOR framework. The purpose of the framework is to provide a tool to support analysis of management and policy options for tropical forests. A simple logging treatment was simulated in the SYMFOR framework by simulating heavy logging, felling an average of 12 trees in per one hectare plot, with the associated damage to the residual stand. It should be

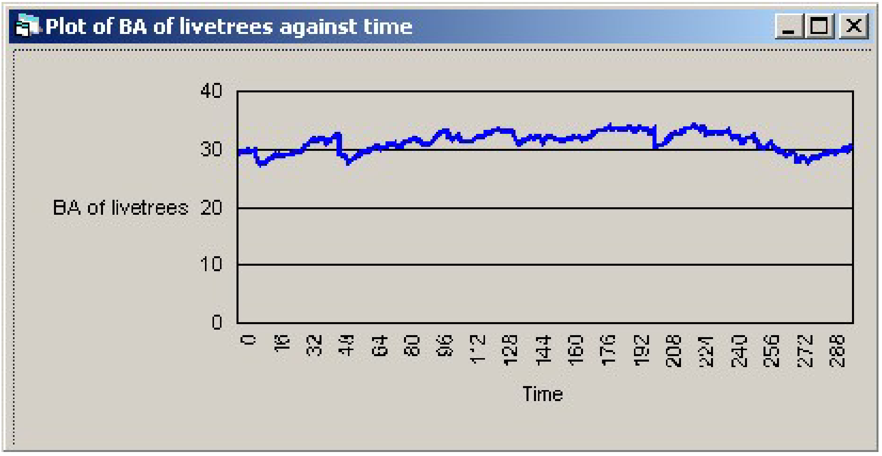


Figure 8: Variation in total basal area with time for 1 ha of unlogged primary forest, in a single simulation. No management interventions were simulated. This figure was taken directly from the time-series plotter in SYMFOR.

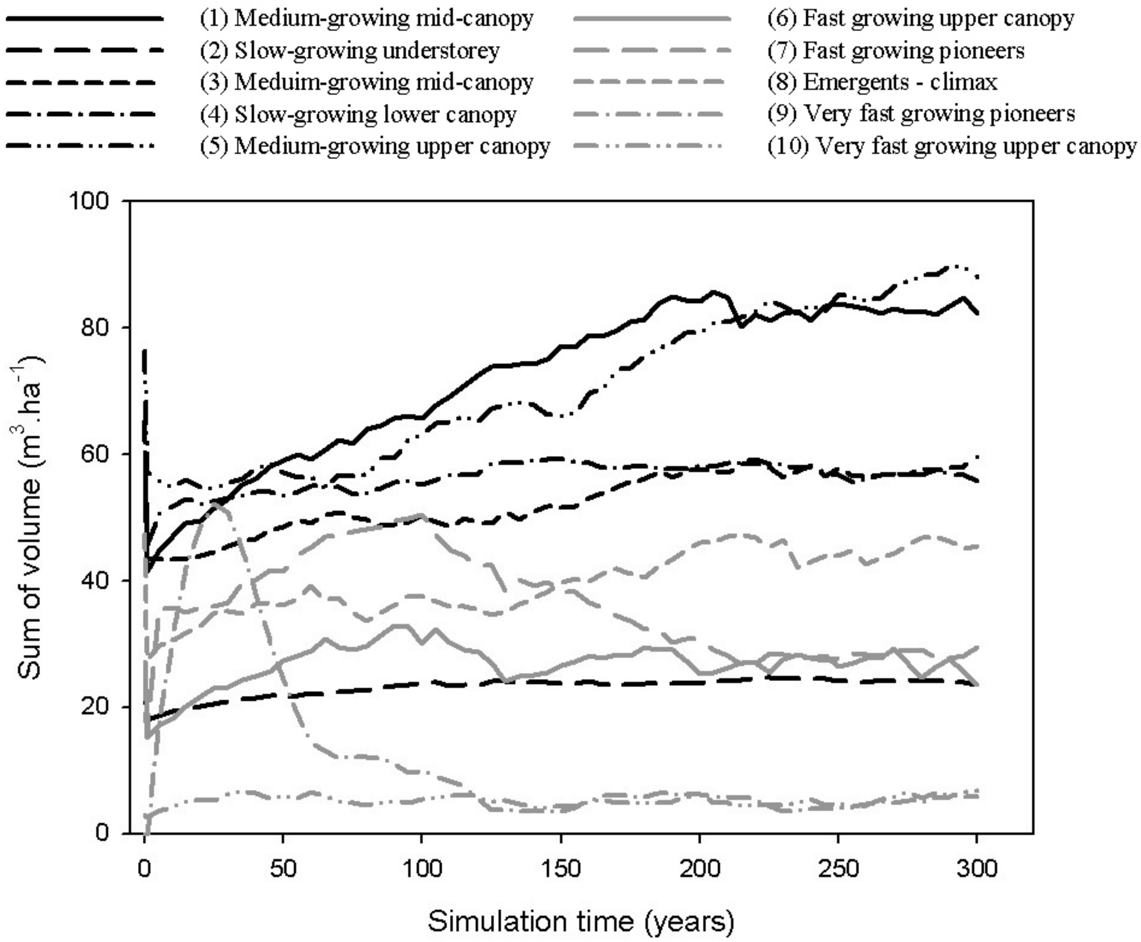


Figure 9: The sum of volume over all trees of each species group against time, showing forest development after logging. The figure shows the mean results over all 10 repetitions of simulations on a 1 ha plot expected to be typical of the region. A dynamic equilibrium, in terms of species composition and size distributions, has been established after approximately 200 years.

noted that this level of extraction is similar to the extreme of current management practice in Brazil and was chosen to demonstrate the ecological response to a severe disturbance. Management practice, the effect of markets and yield regulation not been addressed in this report. The simulation was performed 10 times on a 1 ha plot comprised of 4 typical PSP from unlogged forest in the Tapajos region simulating 300 years of forest development following a single logging operation in the first year.

The simulation predicted that significant changes in species composition should be expected following heavy logging of this forest type. The total volume (Figure 9) and number of stems (Figure 10) of each species group and balance between groups changed significantly following logging. The time required to reach a new equilibrium between the groups is predicted to exceed 200 years, although the total number of trees of each species group stabilises in around 150 years. This is similar to the length of time required for recovery of forests in Southeast Asia (Phillips *et al.*, 2002b), but less than that for inland forests in Guyana (Phillips *et al.*, 2002a), using similar analysis with alternative models in the SYMFOR framework. The reason for the similarities and differences is related to the complex interactions between recruitment, growth, mortality and competition, but is likely to be dominated by the much lower annual growth rates observed in Guyana.

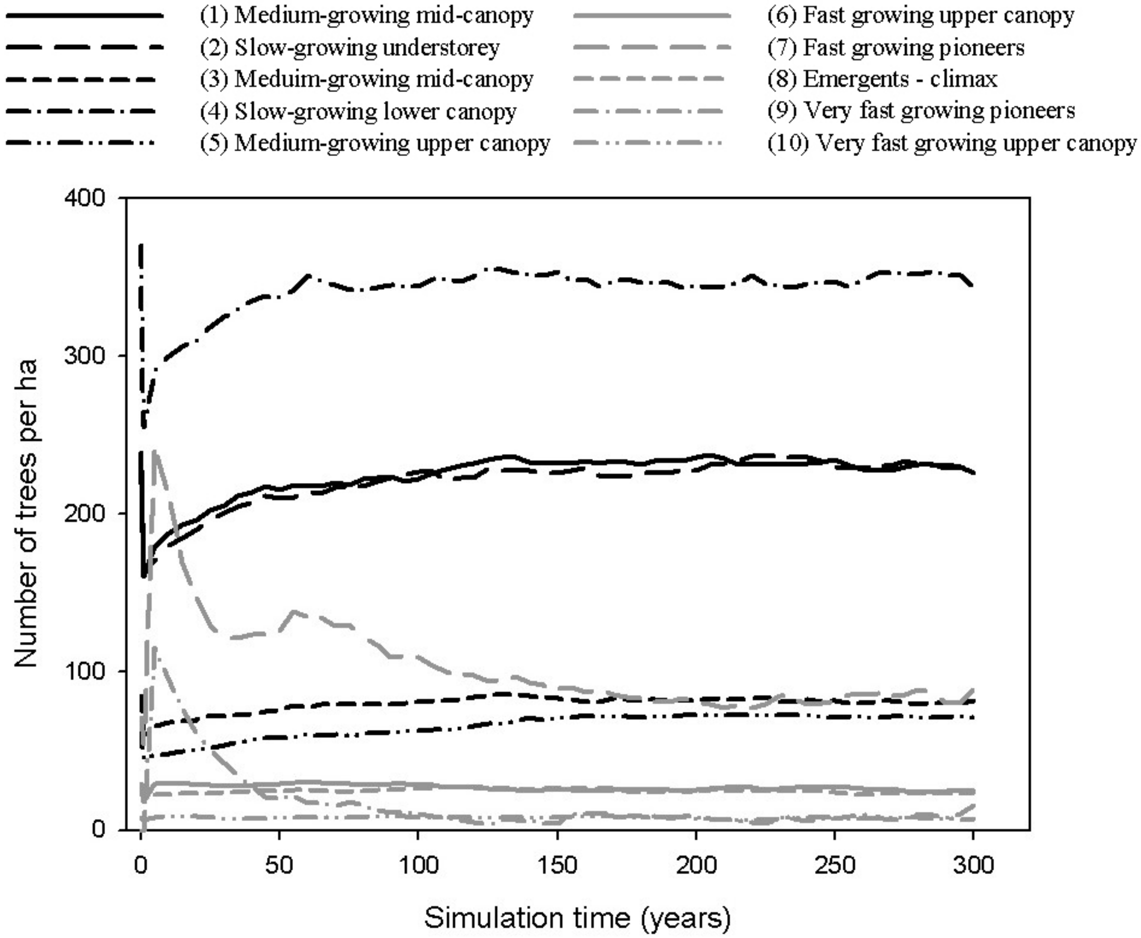


Figure 10: The number of trees of each species group against time, showing forest development after logging. The figure shows the mean results over all 10 repetitions of simulations on a 1 ha plot expected to be typical of the region. A dynamic equilibrium, in terms of species composition, has been established after approximately 150 years.

During these simulations and their subsequent analysis, it was noted that significant differences exist between the plots in terms of species composition. For this reason all simulation studies should be based on multiple simulations using a range of plots.

5 Assumptions and limitations

5.1 Assumptions

This section describes the assumptions required for the development and implementation of the ecological model described in this document. Other models within the SYMFOR framework, including management models, that have their own associated assumptions that are not described here.

For many of the assumptions, their accuracy is not quantifiable, or even meaningful. Even if it were, quantifying it would not quantify the impact of making the assumption on model results (the relationship is not often linear or simple). Their effect on model results contributes to the systematic uncertainty associated with the results, but can be extremely difficult to quantify; while parameter values may be altered to test their effect, this is not usually possible for model assumptions (Phillips and van Gardingen, 2002).

Estimating the systematic error associated with the assumptions made in model design is very difficult, and must be done for each application of the model separately, because it depends on the starting conditions and management model scenarios used, as well as the results that are quoted. Approaches may be used to minimise the effect of such systematic errors, for example by comparisons of two alternative scenarios; the systematic errors will be very similar for the two scenarios, and will largely cancel out. Otherwise, qualitative statements may be used to account for the effects of the assumptions on model results.

Growth model

1. Trees grow as a function of their species group, diameter and competition environment;
2. Trees cannot decrease in stem diameter;

Competition model

3. Trees only compete with larger neighbouring trees;
4. The competition experienced by a tree is not dependent on its location within a grid-square;
5. No competition is experienced by a tree from neighbouring trees that are more than 2 sub-plots away;

Recruitment model

6. The probability of recruitment is a function of predicted growth rate;
7. The probability of recruitment may be non-zero even when the predicted growth rate is zero (since the predicted growth rate represents the mean of a distribution);
8. Seedlings and saplings with a diameter between 0 cm and 5 cm grow with the same diameter increment as trees with a DBH between 5 cm and 6 cm;
9. A constant supply of seedlings exists for all species groups, except shortly after an area has been cleared (as part of a management process);

Mortality model

10. Mortality is stochastic (a random event), and is described as a function of species group and diameter;
11. Trees with DBH less than 80 cm do not fall over when they die, or if they do, then they do not cause damage to other trees in the process;
12. All other trees are assumed to rot whilst standing and cause no additional damage or further competition to the residual stand;
13. The area under which damage may occur resulting from a falling tree is described by a small kite-shape, if lianas have been cut prior to tree-fall, and a large rectangle if they have not;

14. There are no catastrophic mortality events (drought, flood, hurricane, etc) described in the model, and the effects of such processes are captured in the mean rates of mortality;
15. Any mortality due to damage from falling trees or silvicultural activity occurs at the same time as the damage itself, with no time delay;

Other assumptions

16. The behaviour of all trees within a species group is specified in the same way;
17. Trees at the edge of the simulated plot interact with trees on the other side of the plot, through plot-wrapping;
18. All sites are equivalent, in terms of their ability to support tree species compositions and diameter distributions;
19. All sites are equivalent, in terms of their suitability and access for logging and other management activities;
20. Non-modelled phenomena (such as climate, the presence of climbers, pests, disease, watercourses, slope, soil-types, etc.) are constant over time and are the same as for the data used for calibration;
21. The calculations of stem basal area and volume assume that all trees are circular in cross-section;
22. For the calculation of stem volume, all trees have the same relation between diameter and volume;
23. In a given simulated year, the forest processes are simulated in the following order: management, growth, mortality, recruitment.

5.2 Limitations

The model should only be applied to areas of forest with similar soils, species composition and tree density as the plots used to calibrate the Brazil-Tapajos model. The rates of growth, recruitment and mortality should be similar to those observed in the Tapajos and Jari sample plots. In practice, this means primary or managed natural forests in the *Terra Firme* forests of the Eastern Amazon.

The suitability of the model should be evaluated before it applied to new areas or forest types. Data from the new area should be compared with those used for calibration of the Brazil-Tapajos model. The nature of the comparison will depend on whether repeated diameter measurements are available to estimate growth rates.

In all cases the user should compare:

- Stand basal area and volume
- Stand density
- Species composition by individual species and ecological species groups. This comparison should consider the number of stems in each taxa and their total basal area.

If repeated measurements are available from permanent sample plots, these should be used to calculate mean diameter increments for each species groups. These should be compared with the values reported in Table 7.

The method of comparison will need to be subjective and result in an assessment of how similar the new area is to the Tapajos and Jari plots used for calibration. In the absence of sufficient information for the above analysis, a more simple comparison should be made with whatever data is available, and suitable provisos appended to the description of any simulation results.

6 Conclusions

The Brazil-Tapajos ecological model has been developed, calibrated and implemented within the SYMFOR framework. This implementation was then tuned against the criteria that the structure of an unlogged primary forest should not change significantly with time. Validation and model testing were performed by comparison with real data over 16 years following a heavy logging event.

The ecological species groupings and initial simulations with the model describe ecologically relevant characteristics and processes in the forest. The results indicate that this forest is characterised by some very fast growing species, and some very large climax species.

The model is suitable for use with data from primary or managed natural forests in the *Terra Firme* forests of the Eastern Amazon, or other areas of similar forest.

Applications of the model will be published separately.

7 Acknowledgements

The views expressed are those of the authors alone, and not necessarily the institutions that supported the work.

The authors would like to acknowledge the contributions of the Dendrogene Project, Belem, Brazil, and INRA, Kourou, French Guiana, for supporting, commenting on and improving the work described.

The staff of EMBRAPA Amazonia Oriental collected and made available the data for this work, and provided time to discuss it, which were the fundamental prerequisites for the modelling and analysis.

IBAMA and Jari Cellulose Ltd. manage the forest from which the data were collected, and their collaboration in the process is greatly appreciated.

The SYMFOR software, and the previous models developed within the SYMFOR framework, were funded by the United Kingdom Department for International Development (DFID) Forestry Research Programme project R6915 for the benefit of developing countries. This work was carried out by Edinburgh University and partners.

8 References

Alder, D., and Silva, J. N. M., 2000. An empirical cohort model for management of Terra Firme forest in the Brazilian Amazon. Forest Ecology and Management, 130, 141-157.

Degen, B., Gregorius, H. R. and Scholz, F., 1996. Eco-Gene, a model for simulation studies on the spatial and temporal dynamics of genetic structures of tree populations. *Silvae Genetica*, 45, 323-329.

Degen, B. and Scholz, F., 1996. Der Einsatz des Simulationsmodells ÖKO-GEN zur Erarbeitung von Entscheidungshilfen für eine nachhaltige Forstwirtschaftl. In: G. Müller-Starck (ed.) Biodiversität und nachhaltige Forstwirtschaft, 284-299, Ecomed - Verlagsgesellschaft, Landsberg.

Degen, B., Streiff, R., Scholz, F. and Kremer, A., 1997. Analyzing the effects of regeneration regime on genetic diversity and inbreeding in oak populations by use of the simulation model Eco-Gene. In: K. C. Steiner (Ed.), 1997. Diversity and Adaption in Oak Species. College of Agricultural Sciences, Pennsylvania, University Park, Pennsylvania.

Degen, B. and Scholz, F., 1998. Ecological genetics in forest ecosystems under stress as analysed by the simulation model Eco-Gene. *Chemosphere*, 36, 819-824.

Phillips, P.D. 2002a. Joining the simulation models Eco-Gene and SYMFOR. Consultancy report for Dendrogene (DFID), Belem, Brazil. 15th March 2002. 41pp.

Phillips, P.D. 2002b. ExtractSFC: A program to extract complete datasets from the SFC database. Consultancy report for Dendrogene (DFID), Belem, Brazil. 16th May 2002. 23pp.

Phillips, P.D. 2002c. Training in defining ecological species groups within SYMFOR model development. Consultancy report for Dendrogene (DFID), Belem, Brazil. 17th May 2002. 20pp.

Phillips, P.D., van Gardingen, P.R., 2001a. The SYMFOR Framework for Modelling the Effects of Silviculture on the Growth and Yield of Tropical Forests. In the proceedings of IUFRO 4.11 Conference: 'Forest Biometry, Modelling and Information Science', (Editor; K.Rennolls), University of Greenwich, 25-29 June 2001. 12 pages. http://cms1.gre.ac.uk/conferences/iufro/proceedings/

Phillips, P.D. and van Gardingen, P.R., 2001b. The SYMFOR framework for individual-based spatial ecological and silvicultural forest models. SYMFOR Technical Notes Series No. 8, The University of Edinburgh. http://www.symfor.org/technical/framework.pdf

Phillips, P.D., Yasman, I., Brash, T.E., van Gardingen, P.R. 2001. Grouping tree species for analysis of forest data in Kalimantan (Indonesian Borneo). For. Ecol. Man. 157: 205-216.

Phillips, P.D., van der Hout, P., Arets, E.J.M.M., Zagt, R.J., van Gardingen, P.R. 2002a. Modelling the natural forest processes using data from the Tropenbos plots at Pibiri, Guyana. SYMFOR technical note series no. 9, The University of Edinburgh, Edinburgh, 25pp. http://www.symfor.org/technical/index.html

Phillips, P.D., Brash, T.E., Yasman I., Subagyo, S., van Gardingen, P.R. 2002b. An individual-based spatially explicit tree growth model for forests in East Kalimantan (Indonesian Borneo). Ecol. Model. (in press).

Queiroz W.T. de, 1984. Análise de fatores (factor analysis) pelo método da máxima verossimilhança: plicação ao estudo da estrutura de florestas tropicais. Tese de doutor em Agronomia, ESALQ, Universidade de São Paulo, Piracicaba, São Paulo, Brasil.

Silva, J.N.M., de Carvalho, J.O.P, Lopes J.do C.A., de Carvalho M.S.P., 1984. Volume equations for the Tapajos National Forest. Boletim de Pesquisa Florestal Curitiba, No 8/9 p 50-63 Jun/Dez 1984, Embrapa Unidade Regional de Pesquisa Florestal Centro-Sul.

Silva, J.N.M, Araújo, S.M., 1984. Volume equation for small diameter trees in the Tapajos National Forest. Boletim de Pesquisa Florestal Curitiba, No 8/9 p 16-25 Jun/Dez 1984, Embrapa Unidade Regional de Pesquisa Florestal Centro-Sul.

Silva, J. N. M., de Carvalho, J. O. P. d., Lopes, J. d. C. A., Almeida, B. F. d., Costa, D. H. M., De Oliveira, L. C. d., Vanclay, J. K. and Skovsgaard, J. P., 1995. Growth and yield of a tropical rain forest in the Brazilian Amazon 13 years after logging. Forest Ecology and Management, 71, 267-274.

Silva, J.N.M., de Carvalho, J.O.P., Lopes, J.C.A., De Oliveira, R.P., De Oliveira, L.C., 1996. Growth and yield studies in the Tapajós region, Central Brazilian Amazon. Commonwealth Forestry Review 75 (4): 325-329.

Thompson, I, 1997. TREMA - Um Software para Planejamento da Exploração Florestal. International symposium: “Posibilidades de Manejo Forestal Sostenible en América Tropical”, Santa Cruz de la Sierra, Bolivia, 15-20 July 1997. http://bolfor.chemonics.net/Publicaciones/Simposio/thompson.pdf .

van Gardingen, P.R., McLeish, M.J., Phillips, P.D., Dadang Fadilah, Tyrie, G., Yasman, I., 2002. Timber yield, financial and ecological analysis of the sustainable management of logged-over Dipterocarp forests in Indonesian Borneo. For. Ecol. Man. (In press).

9 Appendix A – The model parameters

The following table shows the model parameters as described in the text, a summary of their usage, and the corresponding variable name in the SYMFOR computer program code.

Parameter	Name in SYMFOR	Usage
$z_1 - z_3$	z1coeff, z2coeff, z3coeff	Competition index coefficients
$b_0 - b_2$	dicomp1, dicomp2, dicomp3	Diameter-independent competition index coefficients
$a_0 - a_4$	p0 – p4 (species group)	Growth model coefficients
$r_1 - r_3$	i1, i2, i4 (species group)	Recruitment model coefficients
T_I	ingrowthtime (species group)	The minimum time for a seed to reach a DBH of 5 cm
$m_0 - m_2$	b0, b4, b5 (species group)	Mortality model coefficients
D_{95}	p95 (species group)	An effective maximum tree DBH
b_d	mbinwidth	The width of the smallest DBH bin for mortality modelling
D_m	mindbhfalls	The minimum diameter for tree-fall in natural mortality
P_f	probfalls	The probability of tree-fall in natural mortality
f_C	a (species group)	The fraction of tree height at which the crown begins
H_m	maxheight (species group)	The maximum height of a tree of any diameter
s	startslope (species group)	A coefficient in the equation relating height to DBH