

An individual-based spatially explicit tree growth model for forests in East Kalimantan (Indonesian Borneo)

Paul D. Phillips^{a,*}, T.E. Brash^a, Irsyal Yasman^b, Prapto Subagyo^c,
P.R. van Gardingen^a

^a Centre for the Study of Environmental Change and Sustainability (CECS), The University of Edinburgh, John Muir Building, Kings Buildings, Mayfield Rd., Edinburgh EH9 3JK, UK

^b PT. Inhutani I, Manggala Wanabakti Bld., Block VII, 12th floor, Jl. Gatot Subroto, Senayan, Jakarta Pusat 10270, Indonesia

^c Balai Penelitian Kehutanan Samarinda, Jalan A. Wahab Syahrani, Sempaja, P.O. Box 1206, Samarinda, Kalimantan Timur, Indonesia

Received 1 February 2001; received in revised form 4 February 2002; accepted 4 March 2002

Abstract

A model of ecological processes that determine the behaviour of individual trees, and thus the overall characteristics of forest stands, is described, evaluated and discussed. The purpose of the model is to simulate individual trees as part of the SYMFOR framework for modelling the effects of silviculture on the growth and yield of tropical forests. The model encompasses the response of individual trees to a range of conditions, described in a spatially explicit form, that represent typical situations occurring in forests at all stages following management interventions. The model describes the processes governing the behaviour of trees, allowing it to be used with validity for simulations of management regimes that do not have experimental precedent. The model can thus guide or negate the requirement for lengthy and laborious field trials. Sub-models of tree growth, natural mortality and recruitment processes specify the ecological model. The growth model describes the annual diameter increment for an individual tree, calculated from a deterministic component using tree diameter, local competition (5 m radius), wider competition (30 m radius) and the effect of previous harvesting and a stochastic component representing the remaining variation. The mortality model represents the death of trees as a probability function that increases with diameter. Damage from falling trees is represented by a trapezium-shaped area in which all trees smaller than the falling tree are killed. The recruitment model uses 10 by 10 m² grid-squares to represent the probability of recruitment of trees past the 10 cm diameter threshold of the SYMFOR framework based upon the average competition within a grid-square. The whole model is derived from data and does not rely on theories of tree or forest behaviour, but all aspects represent real forest processes.

© 2002 Elsevier Science B.V. All rights reserved.

Keywords: SYMFOR; Ecosystem; Model; Silviculture; Forest; Tropical; Permanent sample plot; Stochastic; Monte Carlo simulation

1. Introduction

Forestry management and policy in the developing world generally lacks predictions of the future state of forests and their likely responses

* Corresponding author. Tel.: +44-131-650-7861; fax: +44-131-650-7863

E-mail address: paul.phillips@ed.ac.uk (P.D. Phillips).

to management interventions. With this information, it would be possible to make better-informed decisions affecting the people, economy and ecology of forested areas through enhanced sustainable forest management.

For most forested areas in tropical regions, there are limited numerical data describing the impacts of forestry practice that can be used to support the development of sustainable management practices. Anecdotal evidence may be available, but is generally of a subjective nature and site specific. Natural tropical forests are complex ecosystems whose attributes cannot simply be scaled (van Gardingen et al., 1997) in order to obtain accurate predictions of the future state of the forest following management interventions. Forest management tools developed for use in temperate regions, often applied to single aged monocultures, are of little use in mixed age tropical forests characterised by high species diversity. Researchers and forest managers thus require predictive tools that can represent the ecosystem including the underlying processes involved in forest growth, regeneration and management. These tools should be accessible to forest managers and policy makers, allowing them to evaluate aspects of simulated management relating to the forest. The tools should be capable of simulating currently implemented management scenarios as well as new practices that have no direct precedent.

SYMFOR, as described by Phillips and van Gardingen (2001), is a framework that combines models of forest ecology with models describing important elements of management treatments typically applied in selective logging of tropical forests. The SYMFOR framework requires a description of natural forest ecology in the form of an ecological model describing tree growth, mortality and seedling recruitment, and thus the forest as a whole. This study aimed to produce such a model derived from data describing a Dipterocarp forest in Indonesian Borneo (East Kalimantan).

Long-term sustainable management of mixed tropical forests for timber production requires estimates of the growth and yield of the forests for the initial and subsequent harvests. This is

difficult because of the complex structure and species composition of these forests and because most existing management records only describe production from the first harvest from primary forest. The implementation of sustainable management may require predictions of quantities in addition to the projected timber yield as a function of time. The development of multiple objective forest management may also consider production of non-timber forest products and maintenance of biodiversity. These aspects require models that can capture the ecology of the forest stand in terms of species composition, size classes and if possible the spatial distribution of individual trees.

Models that have already been developed describing the ecology, growth and yield of tropical forests include FORMIX (Huth et al., 1997, 1998) and elements of DIPSIM (Kleine and Heuvelodp, 1993; Ong and Kleine, 1996). These are not suitable for application in this study since they are not based on individual trees or spatially explicit. FORMIND, developed in parallel with SYMFOR by Köhler and Huth (1998), is individual-based but not spatially explicit: it represents the number of trees of a given species group in a size-class by a single tree, that is grown as an individual. The approach used by Gourlet-Fleury and Houllier (2000) for a model of forest growth in French Guiana is more applicable to the current study. Individual trees grow in competition with one another within a plot, and mortality and seedling recruitment are represented as separate processes. This is the approach we have used to develop an ecological model for SYMFOR for Indonesia.

1.1. Overview of the SYMFOR framework

The SYMFOR framework provides a WindowsTM-based interface for the user to:

- select input data, a pre-defined model and output data requirements;
- run the simulation and view graphical displays of the data;
- edit either the ecological or silvicultural models: select different modules for a particular func-

tion, or change the values of the module parameters.

Fig. 1 shows an ecological model within the SYMFOR framework. During a simulation, the SYMFOR framework calls functions for growth, mortality and recruitment each once per simulated year. The modules that satisfy these functions are selected individually by the user or as part of an entire saved model that the user can select. Within SYMFOR there are three separate sections to the ecological model that require development, description and representation in the framework:

- tree growth (diameter increment);
- natural tree mortality; and
- new tree recruitment (the growth of saplings past a diameter of 10 cm, representing the growth of new trees into the model).

Data that are input to the SYMFOR framework by the user are based on Permanent Sample Plot (PSP) data (Alder and Synott, 1992). The important elements of these data for modelling purposes are descriptions of the position, diameter and species for each tree in the plot. In addition, the user specifies when and how heavily the plot was last logged as the number of years since logging

and the number of live stems remaining in the stand following logging.

The SYMFOR framework utilises species grouping in order to facilitate the implementation of the model, allow a realistic representation of the less frequent species occurring in the forest and to facilitate concise descriptions of the forest state to be made in output summaries. An ecological species grouping described by Phillips et al. (2002) has been implemented in the SYMFOR framework for the ecological model in Indonesia, and is summarised in Appendix B. A separate commercial species grouping is implemented for the silvicultural component of the model (Phillips and van Gardingen, 2001).

2. Model description

The model is described with three components: (1) a tree growth model, (2) a tree mortality model and (3) a recruitment model. In the following text, all probabilities and random numbers used in the model are within the range 0.0–1.0 unless otherwise stated. Where parameters are evaluated by regression against real data, the data used are from all plots and all re-measurements unless otherwise stated.

A legend to the symbols used in the text and mathematical functions is included in Appendix A.

2.1. Growth model

2.1.1. Diameter increment model

Annual diameter increment (ΔD) for each tree is calculated from two terms:

$$\Delta D = \Delta D_d + \Delta D_s \quad (1)$$

where ΔD_d is the deterministic component of growth and ΔD_s is a stochastic component; a term that describes the variation remaining in the data that was not explained by the deterministic term. ΔD_s is calculated using Monte Carlo techniques (Fishman, 1995).

2.1.2. Deterministic component

The deterministic growth term arises from regression of the growth data against several

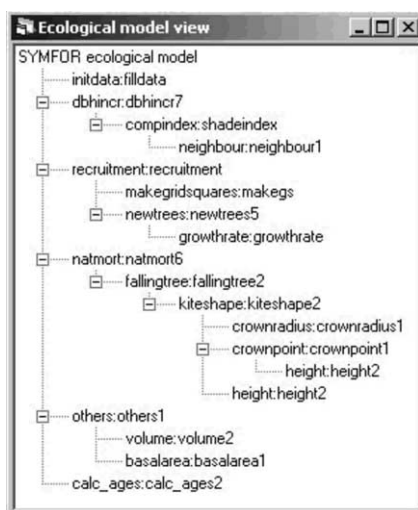


Fig. 1. An example of the structure of an ecological model within the SYMFOR framework.

dependant variables. An index of competition, known as ‘shade-index’, based on the relative size and distance between a tree and its neighbours was developed. It was used to form two dependant variables; with neighbouring trees defined as those within a radius of 5 m (‘shade-index at 5 m’, S) and 30 m (‘shade-index at 30 m’, S') of the subject tree. It is calculated for the subject tree, i , by the equation

$$S_i = \sum_{j=1}^n \frac{D_j}{D_i d_{ij}} \quad (2)$$

where j indicates a tree identified as a neighbour, n is the total number of neighbours identified, and D_i is the diameter at breast height (DBH) for the subject tree, D_j is the DBH for the neighbouring tree and d_{ij} is the distance between the subject tree and its neighbour. Only living trees are considered in this calculation as trees that have died but remain standing are not considered to compete with the live trees. S' is calculated using the same equation for all competing trees within a 30 m radius of the subject tree. SYMFOR implements plot wrapping at the border to remove edge effects when calculating the competition index (Phillips and van Gardingen, 2001).

The deterministic part of Eq. (2) can be expressed as four terms corresponding to the dependant variables:

$$\Delta D_d = G_d G_S G_{S'} + G_Y \quad (3)$$

where G_d , G_S , $G_{S'}$ and G_Y are the growth terms corresponding to, respectively, the dependent variables DBH, shade-index at 5 m, shade-index at 30 m and previous logging intensity.

The DBH term, G_d , in Eq. (3) is given by:

$$G_d = (p_0 + p_1 D) \exp(-p_2 D^2) + p_3 \quad (4)$$

where D is the tree DBH minus 10 cm and p_0 , p_1 , p_2 and p_3 are parameters to be evaluated by regression. The form of this equation is such that for small trees the term tends towards $p_3 + p_0$, and at very large tree diameters it tends towards p_3 .

The shade-index at 30 m term, $G_{S'}$, in Eq. (3) is given by:

$$G_{S'} = \exp(-p_4 S'^2) \quad (5)$$

where p_4 is a parameter to be evaluated by regression. The form of this equation is such that for low competition the term tends towards unity, and at very high competition it tends towards zero.

The shade-index at 5 m term, G_S , in Eq. (3) is given by:

$$G_S = 1 + p_5 S \quad (6)$$

where p_5 is a parameter to be evaluated by regression.

The term in Eq. (3) that represents the variation in growth rates with previous logging intensity, G_Y , depends on the number of years since logging, and the logging intensity at that time:

$$G_Y = p_6 Y \exp(-p_7 Y) F_Y \quad (7)$$

where p_6 and p_7 are parameters to be evaluated by the regression, Y is the number of years since logging and F_Y is a scaling term. The scaling term is given by:

$$F_Y = 3 \left(1 - \frac{N_T}{p_8} \right) \quad \text{if } N_T \leq p_8 \quad \text{and}$$

$$F_Y = 0 \quad \text{if } N_T > p_8 \quad (8)$$

where p_8 represents the average number of trees in a plot, against which logging intensity may be estimated. This term is necessary to represent the intensity of the logging activity, which is represented by the number of trees remaining in the stand immediately following logging, N_T . The factor F_Y is designed to be very low for low logging intensity and high for heavily logged stands.

2.1.3. Stochastic component

The variation of growth rates between trees that is not explained by the deterministic component of growth is described by the histogram of the residuals, r , where:

$$r = \Delta D_O - \Delta D_d \quad (9)$$

and ΔD_O is the observed annual growth of a given tree. The histogram represents a distribution of the probability of any given tree having a particular value, g_r , of r . This distribution, $P(g_r)$, is modelled with a Gaussian function:

$$P(g_r) = \exp\left\{-\frac{(g_r - p_{12})^2}{2p_{11}^2}\right\} \quad (10)$$

where p_{11} and p_{12} are parameters to be evaluated by the regression of the residuals against the probability distribution. p_{11} represents the width of the distribution and p_{12} represents the value of g_r at the peak. This distribution is sampled for possible values of ΔD_S , g_r , within the range:

$$p_{12} - 4p_{11} \leq g_r \leq p_{12} + 4p_{11}. \quad (11)$$

In order to generate the value for ΔD_S for a given tree, a candidate value, g_r , is selected using a random number within the range specified in Eq. (11). $P(g_r)$ is calculated for the candidate value, g_r , and a second random number, r_1 , is chosen to decide whether the candidate is accepted; r_1 is chosen such that it covers the full range of $P(g_r)$. If $r_1 < P(g_r)$:

$$\Delta D_S = g_r \quad (12)$$

and otherwise a new value of g_r is selected and the test repeated. Monte Carlo sampling (Fishman, 1995) in this manner is known as an acceptance/rejection method.

Serial correlation is included in the process of sampling for a candidate value of ΔD_S . Two random numbers are chosen: one, r_2 , to represent the serial correlation, the other, r_3 , to represent the remaining random component. The value of the serial correlation random number, r_2 , is fixed for a given tree for the whole simulation, and is selected when the tree first appears in the model; either input as initialisation data or when created by the recruitment model. The value of r_3 is chosen each time growth is calculated for a given tree. These numbers are combined to produce g_r , the candidate value of ΔD_S , as follows:

$$g_r = [r_1 p_{10} + r_2(1 - p_{10})]4p_{11} - 4p_{11} + p_{12}. \quad (13)$$

The probability of the value of g_r being accepted as the value of ΔD_S is given by $P(g_r)$ in Eq. (10), assuming a uniform distribution of g_r . But the frequency distribution of g_r is not uniform because of the process of adding two random numbers together; this gives a ‘hill-top’ distribution. The distribution, given by $Q(g_r)$, is expressed by:

$$Q(g_r) = \left\{ \begin{array}{ll} g_r < g_1, & Q_m + \frac{1 - Q_m}{g_1 - g_m}(g_r - g_m) \\ g_1 \leq g_r \leq 1 - g_1, & 1 \\ g_r > 1 - g_1, & 1 - \frac{1 - Q_m}{g_1 - g_m}(g_r + g_1 - 1) \end{array} \right\} \quad (14)$$

where:

$$g_1 = \min\left[\frac{p_{10}}{1 - p_{10}} \right], \quad (15)$$

Q_m is an introduced minimum value of $Q(g_r)$ (see below) and

$$g_m = p_{12} - 4p_{11} \quad (16)$$

(the minimum of the range of possible values of g_r).

The distribution of ΔD_S is not uniform, and for this reason the probability distribution, $P(g_r)$, must be corrected to obtain $P'(g_r)$, the probability distribution for the actual distribution of g_r , $Q(g_r)$:

$$P'(g_r) = \frac{P(g_r)}{Q(g_r)} \quad (17)$$

$P'(g_r)$ is then normalised so that it reaches a peak at a value of 1.0.

2.2. Recruitment model

Recruitment was defined as depending on two criteria for each year of the simulation: the likely growth rate of a tree at a particular location, and the species group of the tree.

The stand area in SYMFOR is divided into a mesh of grid-squares for the purpose of representing the recruitment process. Within each 10×10 m grid-square a sample location is chosen for each species group. A hypothetical new ‘virtual’ tree is created at that location, and the deterministic component of its growth rate is evaluated (ΔD_d , Eq. (3)) and converted to a recruitment probability, P_i , for species group i , using the equation:

$$P_i = m_i \Delta D_d + c_i \quad (18)$$

where m_i and c_i are coefficients evaluated by regression.

The presence of disturbed areas following harvesting will influence the rate of recruitment.

Recruitment will not happen inside these areas until the time taken for trees to grow from seed to a DBH of 10 cm is exceeded. This minimum time is represented by the parameter t_i , the ‘in-growth time’, for each species group, and is doubled for skid-trail areas to represent the inhibition of growth by soil compaction. The position of the simulated new tree is tested to determine whether it is inside a disturbed area and, if so, the in-growth time is compared with the time since each clearance occurred. The tree will only become established if the number of years since the most recent disturbance was sufficiently long enough to permit re-growth of a tree of the species group at that position. Otherwise another position within the grid-square is randomly generated and tested as above. In the unlikely event that the whole grid-square has been recently disturbed, the model abandons recruitment in that grid-square after 100 attempts and moves on to the next.

2.3. Model of natural mortality and associated damage

The probability of mortality of individual trees P , is modelled on an annual basis as a function of DBH, D . This represents a linear increase in mortality probability with tree diameter up to a threshold diameter D_m , above which mortality is increased by a factor of ten. The probability is then given by:

$$P = \left\{ \begin{array}{ll} 10 \leq D \leq D_m, & m_1 + m_2 D \\ D_m < D, & 10(m_1 + m_2 D) \end{array} \right\} \quad (19)$$

where m_1 and m_2 are parameters that have a value for each species group. The probability of mortality is calculated once for each year, each living tree, and compared with a random number, r_4 . The tree will die if $r_4 < P$. This is represented with the SYMFOR framework by converting the ‘live-tree’ to a ‘fallentree’ software object.

The mortality model interacts with the serial correlation term (r_2) of the modelled growth rates of large trees (Eq. (13)). Slower growing trees tend to die before they become large, leaving the large trees to have a mean r_2 greater than 0.5. This is an artefact of the calibration because the assumption was made that the mean r_2 was 0.5 for all values of

D . Consequently, the mean modelled growth rate of large trees is larger than the deterministic prediction that was calculated from the data. To restore the correct average values, r_2 is reduced by 0.0017 for each cm that the tree grows.

The probability of a dying tree falling over is represented by the parameter m_3 which is compared with random number, r_5 . If $r_5 < m_3$, the tree will fall, and the damage resulting from the tree falling is calculated. The direction of tree-fall, α_f , is assigned using the equation:

$$\alpha_f = 2\pi r_6 \quad (20)$$

where r_6 is a random number and α_f is in radians measured as a bearing from the $+y$ axis of the plot co-ordinate system.

2.3.1. Damage resulting from falling trees

A region in the plot affected by a tree falling is created as a trapezium-shaped area, with a width equal to twice the crown-radius, C_R , of the falling tree, and a length of the height, H , of the falling tree. The widest point is situated at a distance equal to the crown-height, C_p , from the base of the tree. These parameters are generated from relationships with DBH defined within the SYMFOR framework for each species group (Phillips and van Gardingen, 2001).

All trees within the area defined as under influence from a falling tree are examined as candidates for suffering damage. Only candidate trees that are smaller than the falling tree are defined to be killed by damage, and are converted from ‘livetree’ software objects to ‘smashedtree’ software objects. The assumption is made that there is no secondary damage, so the effects of the new ‘smashedtree’s falling over is not calculated. Similarly, the effect of trees that died from natural mortality but did not fall over have no further effect on the simulated forest.

3. Data and parameter estimation

3.1. Data

The data used for development of this model are permanent sample plot data from the Berau region

of East Kalimantan (Indonesian Borneo). Lowland mixed dipterocarp forest is predominant in this region. The area is managed natural forest, some parts of which have been selectively logged. The plots were set up by the forestry company PT. Inhutani 1 as part of the STREK project (Bertault and Kadir, 1998) and have been maintained with the help of the European Union funded Berau Forest Management Project since 1996.

A detailed account of the history of the plots, the treatments applied and the data collection is given by Bertault and Kadir (1998), and we have kept this description to a brief summary. The region is at a latitude of 1–2°N and low altitude. The terrain is locally hilly, but generally in a shallow watershed, experiences a rainfall of 1800–3000 mm per year and low temperature fluctuations between 25 and 35 °C. The dry and wet seasons are not extreme, and the relative humidity remains high throughout the year.

There are 72 permanent plots, each a square 1 ha arranged with four adjacent plots arranged in square blocks of 4 ha. Data from four surveys of the plots were available at the time of analysis, with approximate dates for the surveys being 1991, 1993, 1995 and 1997. 24 ha of forest were logged in 1980; all other plots were primary forest until experimental treatments were carried out in 1992.

Of the 48 ha used for trial silvicultural treatments:

- 12 ha remained as unlogged control plots;
- 12 ha were logged according to the conventional selective logging system as implemented in Indonesia at that time, where all commercial stems with diameter greater than 50 cm were potentially harvested;
- 12 ha were logged using a Reduced Impact Logging (RIL) treatment (Bertault and Sist, 1997; Sist et al., 1998; Sist and Bertault, 1998; Dykstra and Heinrich, 1996; Elias, 1998) with a cutting diameter lower limit of 50 cm;
- and 12 ha were logged using RIL with a cutting diameter lower limit of 60 cm.

The data available for each tree in each plot and each survey were DBH, a numerical code representing species and the co-ordinates of the trees'

position relative to an origin at one corner of each 1 ha plot. Data for all living trees with a DBH greater than 10 cm were recorded. In-growth (recruitment) and mortality were recorded at each measurement campaign. The data were collated within a database system (Rombouts, 1997) that was used to validate and where possible correct data before use for model calibration. This system was then used to identify the remaining trees with unreliable diameter measurements that were excluded from subsequent analysis, with the exception of the analysis of competition.

3.2. Parameter estimation

3.2.1. Growth model

The growth model was described as equations Eqs. (1)–(17). The parameters p_0 – p_7 for the deterministic component of the model were evaluated using non-linear regression techniques separately for each species group. Parameter p_8 represents the number of trees in unlogged primary forest and this was set to the mean value of 530 stems per ha estimated from the unlogged STREK plots.

The remaining parameters for the growth model described the stochastic component of growth (Eqs. (9)–(17)). The hypothesis was tested that serial correlation exists between residual values; that is, that variation in diameter increment is correlated in time (Vanclay, 1994). Pairs of sequential diameter increment measurements were used to test the correlation with time. It was found that for every species group approximately 75% of pairs of sequential diameter increment measurements ($N_{a,b}$) were either above or below the predicted value, compared with the 50% (N_e) expected for a truly random distribution. The serial correlation, p_{10} , was calculated as:

$$p_{10} = 2(N_{a,b} - N_e) \quad (21)$$

Thus a serial correlation of approximately 50% was observed.

Within the range specified in Eq. (14), the Gaussian function was shown to fit the histogram of residuals with an R^2 value of more than 90% for all species groups except group 5, which had fewer

observations with which to construct the histogram.

All parameters, p_{10} – p_{12} , were evaluated separately, as described above, for each of the species groups. Their values are given in Table 1.

3.2.2. Recruitment model

The parameters m_i and c_i (Eq. (18)) are coefficients evaluated by regression using data from all plots. Growth rates of trees observed in one survey that were not present in previous surveys were calculated using Eq. (3). The growth rate of a hypothetical tree of each species was evaluated for a random location in each grid-square. The data were collated by growth rate, for each species group, to calculate the probability of recruitment for a given growth rate and species group. The probability was modelled as a linear function of diameter.

The minimum ingrowth time, t_i , was evaluated from typical growth rates of each species group at a DBH of 10 cm.

All recruitment parameter values are given in Table 2.

3.2.3. Model of natural mortality and associated damage

The mortality functions were defined as Eq. (19). The probability of a given tree dying in a

Table 2

The parameters m_i and c_i associated with the probability of recruitment per grid-square per year for each of the ten species groups, i (labelled 1–10)

i	m_i	c_i	t_i
1	0.0113	−0.0013	8
2	0.0484	0.0	20
3	0.0606	0.0	15
4	0.0656	0.0	18
5	0.0734	−0.0291	5
6	0.3518	−0.1366	8
7	0.1530	0.0	12
8	0.5000	0.0	18
9	0.0663	0.0	20
10	–	–	–

The parameters t_i , the time taken (years) for seedlings of species group i to reach a DBH of 10 cm, are also shown. Note that there is no recruitment into species group 10.

particular year from natural causes depends on the species group. No significant correlations were found in the data between the probability of death and tree diameter, previous growth rate, basal area or any competition index for an individual species group. The lack of statistically significant correlation between mortality probability and any available variable was probably a result of the small sample of tree mortality in the available data. Only plots that were unlogged for the duration of

Table 1

The parameters p_0 to p_{14} used in modelling individual tree growth and their values for each of the ten species groups

	Scale	1	2	3	4	5	6	7	8	9	10
p_0	$\times 1$	14.2	2.5	6.1	2.2	197	18.4	2.4	2.7	1.7	6.6
p_1	$\times 10^{-1}$	1.27	5.0	5.0	1.55	0	5.56	3.82	0	0	1.45
p_2	$\times 10^{-1}$	0.87	1.61	1.41	1.43	3.85	1.07	2.16	1.07	1.18	2.77
p_3	$\times 10^{-1}$	6.39	5.19	7.16	5.66	89.9	0	5.53	4.57	5.28	6.63
p_4	$\times 10^{-1}$	2.58	1.98	2.16	1.74	5.71	3.34	1.83	1.86	1.65	2.30
p_5	$\times 1$	0.07	0.15	0.07	0.14	1.31	0.25	0.08	0.12	0.11	0.15
p_6	$\times 10^{-2}$	4.14	6.17	3.41	3.47	0	11.6	2.38	1.50	3.63	2.87
p_7	$\times 10^{-2}$	5.31	9.62	5.78	9.95	14.2	10.6	10.0	7.19	20.7	15.2
p_8	$\times 1$	530	530	530	530	530	530	530	530	530	530
p_{10}	$\times 1$	0.43	0.38	0.43	0.39	0.48	0.48	0.35	0.36	0.33	0.43
p_{11}	$\times 1$	0.48	0.18	0.28	0.14	0.47	0.47	0.14	0.12	0.10	0.15
p_{12}	$\times -10^{-2}$	5.91	2.56	5.69	2.83	9.27	9.27	3.90	3.17	2.15	4.71

The scale is a multiplier for all ten values of a parameter (i.e. the scale multiplied by the value in the Table gives the true parameter value).

monitoring were used, in order to separate the effects of management interventions from the process of natural mortality. Only 900 instances of mortality were recorded from the 20 ha of unlogged plots over the 6 year recording period. These data were unevenly distributed between species and diameter, making calibration of the mortality function unreliable.

An operational mortality model was developed using the assumption that an unlogged forest plot should remain in dynamic equilibrium. In practice, this required an assessment of the average number of trees in the plots and the distribution of stem diameters. The STREK plots had an average of 530 living trees per ha and this, combined with the requirement to maintain a constant diameter frequency curve, was used to deduce the mortality rate. The complex interactions between the growth and recruitment models made it impossible to calculate mortality functions analytically. Average mortality probabilities and maximum diameters for each species group were calculated using mortality data. Trial simulations showed that the probability of mortality should increase slowly with diameter to maintain the diameter frequency curve. This was in contrast to other studies (Clark and Clark, 1992; Newbery et al., 1999) because they included mortality arising from damage in the definition of natural mortality, thus including the mortality of large numbers of small trees. Increased mortality probability at a threshold diameter was necessary to prevent occurrences of trees with unrealistically high DBH values.

Eq. (19) includes the parameters m_1 and m_2 that have a value for each species group. D_m was evaluated from the data by considering the minimum diameter of the largest 1% of individual trees. The values of parameters m_1 , m_2 and D_m are given in Table 3.

While the data used for model development did not include information regarding the cause of death of trees, there was one descriptor indicating whether the dead tree remained standing or fell over. These data were analysed to calculate the probability of a dead tree falling over in the same year that it dies, m_3 . This parameter was calculated as 0.44 for all species groups.

4. Model precision and systematic errors

Evaluation of a model includes the estimations of precision of results, an appraisal of the range of model validity and tests of model performance (Vanclay and Skovsgaard, 1997). The ecological model makes predictions relating to growth and mortality of individual trees in a forest stand. Users are, however, more interested in aggregated statistical summaries of these data, such as the total number of stems in a species group, and the sum of their total basal area. This means that the model should be evaluated on the basis of these aggregated result variables rather than the predictions for individual trees. For this reason a set of standard result variables was used in the evaluation of the performance of the model, its precision and validity. This approach permits users to assess the accuracy of the results of typical simulations from the model.

4.1. Definition of standard result variables

Three sets of result variables were generated for the evaluation of the ecological model described in this paper. These were:

- Number of trees.
- Sum of tree basal area.
- Sum of volume of commercial trees. (Commercial species are defined in the SYMFOR framework (Phillips and van Gardingen, 2001) and must have a diameter greater than 50 cm and a stem quality value greater than 0.3).

Four hundred and fifty result variables were generated in each result set by collating data into the following classes,

- ten species groups;
- four size classes (10–30, 30–50, 50–70, > 70 cm), plus an additional result variable containing all size trees; and
- nine time points after the simulation began (5, 10, 20, 30, 35, 40, 50, 75 and 100 years).

The choice of these result variables is arbitrary but based on typical results used by current users of the SYMFOR framework. It should be noted

Table 3

The parameters m_1 , m_2 and D_m associated with the probability of natural mortality for each of the ten species groups (labelled 1–10)

	Scale	1	2	3	4	5	6	7	8	9	10
m_1	$\times 10^{-3}$	16	6	7	5	60	60	0	0	3	8
m_2	$\times 10^{-2}$	3.0	0	0.5	2.8	8.0	8.0	7.0	5.0	2.0	1.1
D_m	$\times 1$	110	120	120	90	50	50	75	80	65	80

that there is an overlap in the size classes since one result variable contains all size trees. The time intervals are uneven to cover periods of interest in the application of the model. The 35 year interval is included to represent the length of the standard cutting cycle in Indonesia (Departemen Kehutanan, 1993).

Many of the result sets for the sum of commercial volume contained no trees because the commercial definition utilised a definition of a minimum diameter of 50 cm and ecological groups 6–10 did not include commercial species.

4.2. Evaluation of systematic errors

Systematic error describes the uncertainty in the value of a result obtained from a model that arises from the design or calibration processes. Model development should include the estimation of the systematic errors so that users have an understanding of the level of precision or uncertainty of model predictions. Sources include the effect of assumptions made in designing the model, and the effect of the uncertainty in the values of the parameters. Systematic errors for complex models are frequently omitted from discussion (Gourlet-Fleury and Houllier, 2000; Huth et al., 1997; Kirschbaum, 2000; Köhler and Huth, 1998; MacFarlane et al., 2000; Valentine et al., 1998) because the process of evaluating them is too difficult, lengthy or itself error-prone. In this section we describe the evaluation of systematic errors for the model described in this paper.

4.2.1. Method

A common technique used to evaluate the effect of the uncertainty in parameter values on model

results is to vary the value of all or selected parameters (Vanclay and Skovsgaard, 1997). The new results are compared with control results (with all parameters at their original values), and any change in result is an estimate of the systematic error arising from the uncertainty in that parameter value. This technique is similar to more rigorous sensitivity analysis, the differences being that the only parameter values are varied (not initial data), and the amount that the parameters are varied is set to be equal to the parameter uncertainty rather than to the range of possible values.

Complex models such as the ecological model described in this paper may contain many individual parameters making comprehensive sensitivity analysis very time consuming. The ecological model in SYMFOR contains 212 parameter values. In such cases it is impracticable to vary every parameter separately to estimate the systematic errors for each result variable. Software exists to perform sensitivity analyses on complex models (Saltelli et al., 1999), but it is not suitable for analysis using result variables as described above without making very significant changes to the structure of the model and its internal representation of data. Such approaches were thus not appropriate for the evaluation of the SYMFOR framework because of the wide range of results of interest to users. The approach described here is similar to that applied to the evaluation of the FORMIND model (Köhler et al., 2001)

4.2.2. Test parameter sets

The evaluation of systematic errors estimated the effect of varying the value of selected individual parameters on the standard result variables

described above. Each parameter value was changed by an amount representing its uncertainty. Stochastic variation in the results between individual model simulations was removed by calculating the mean from 30 simulations for each parameter variation. These were compared with those obtained from the mean of 100 control simulations where all parameters were set to the original values obtained during the calibration process.

Twenty one test sets were generated where the value of one selected model parameter was varied based on the following criteria for those that were most likely to influence results obtained from the model:

- Most populous species groups.
- Species groups with the largest trees, since large trees often dominate the forest dynamics.
- Parameters from each sub-model that were evaluated from the data.
- An additional ‘dummy’ parameter was created to describe the uncertainty in the deterministic growth model. This was required because statistical analysis indicated significant correlation between parameters for this component of the model.

The uncertainty in each selected parameter value was estimated from either statistical analysis (standard error) or expert knowledge. The amount by which each parameter was varied was double the estimate of the uncertainty on that parameter value. Table 4 shows the parameters, their purpose, their values and the amount and direction of their variation as used in this analysis.

4.3. Calculation of systematic errors

The calculation of systematic errors for stochastic models needs to allow for a random or stochastic error term. The random error is independent of the systematic error (Barlow, 1989), so the total error, e_T , is given by:

$$e_T^2 = e_R^2 + e_S^2 \quad (22)$$

where e_R is the random error term and e_S is the total systematic error term.

Stochastic models must be run several times for each set of parameters, j , such that a mean value, $t_{i,j}$, for each result variable, i , and error on the mean, $e_{i,j}$, are calculable. The control mean value, $t_{i,C}$, and error, $e_{i,C}$, for result variable i , from the parameter set containing the calibrated parameter values with no alterations, can then be compared with each test result in turn in order to calculate the difference, $T_{i,j}$:

$$T_{i,j} = t_{i,j} - t_{i,C} \quad (23)$$

with an associated error, $\sigma_{T_{i,j}}$:

$$e_{T_{i,j}}^2 = e_{i,j}^2 + e_{i,C}^2 \quad (24)$$

For a particular result i , the difference, $T_{i,j}$, is expected to be normally distributed with a mean of zero. Significant systematic errors in the model can thus be tested against critical values. The difference $T_{i,j}$ is normalised by the error term $e_{T_{i,j}}$ to give a proportion $F_{i,j}$ (Eq. (25)).

$$F_{i,j} = \frac{T_{i,j}}{e_{T_{i,j}}} \quad (25)$$

A decision of what value of $F_{i,j}$ is considered to be significant depends on the number of result variables being included. There are many result variables, however, and the limit for a single result must be adjusted to account for the presence of n results:

$$C_1 = \sqrt{C} \quad (26)$$

where C is the confidence limit for any one of n results (with a value of 0.95 or 0.99).

Assuming a Gaussian distribution for results without significant systematic errors, the confidence limit C_1 can be equated to a critical proportion, p_c , of the known statistical uncertainty, $\sigma_{T_{i,j}}$, using tables such as those in Barlow (1989). The results that differ significantly from zero may be identified:

$$|F_{i,j}| > p_c \quad (27)$$

The analysis in this study used only result sets based on more than five individual trees producing a total of 3528 comparisons. The 99% confidence limit for 3528 results gives $C_1 = 0.9999$ (Eq. (26)). Tables to convert a confidence limit into a

Table 4

The parameters that were varied during the evaluation of systematic errors for the SYMFOR model (Phillips and van Gardingen, 2001), their meanings, default values and test values used as the variations

<i>j</i>	Module	Parameter	Species group	Value	Varied value
1	Height	H_m	All	45	55
2	Height	s	All	200	250
3	Crown-point	f_c	All	0.55	0.7
4	Crown-radius	f_r	All	0.25	0.3
5	Growth (deterministic)	Multiplier	2	1.0	1.05
6	Growth (deterministic)	Multiplier	3	1.0	1.05
7	Growth (deterministic)	Multiplier	8	1.0	1.02
8	Growth (stochastic)	p_{10}	3	0.43	0.6
9	Growth (stochastic)	p_{11}	3	0.28	0.30
10	Growth (stochastic)	p_{11}	8	0.12	0.13
11	Mortality	m_1	8	0.0	0.001
12	Mortality	m_2	2	0.0	0.005
13	Mortality	m_3	All	0.44	0.6
14	Mortality	D_m	2	120	140
15	Mortality	D_m	3	120	140
16	Mortality	D_m	8	80	90
17	Recruitment	t	2	20	25
18	Recruitment	t	3	15	20
19	Recruitment	m	8	0.5	0.525
20	Recruitment	m	2	0.0484	0.0509
21	Recruitment	c	3	0.0	-0.002

proportion of the uncertainty often do not extend so close to unity, but may be calculated. A C_1 value of 0.9999 was calculated to correspond to a p_c value of between 4 and 5. The value of 4.0 was taken as a conservative estimate of the 99% limit of confidence, p_c , for any one result.

The data used for the evaluation of systematic errors were from the same set as was used for model development. The model was initialised with data from a plot that had been logged over 1 year before data collection. This was intended to give maximum variation between time-points that could be affected by the parameter variations.

4.3.1. Analysis of significant systematic errors

Analysis indicated 34 individual result sets that had one of more values of $F_{i,j}$ exceeding the critical value of 4 (Table 5) coming from only seven of the 21 test parameter sets (Table 4).

Test set eight modified the serial correlation of the stochastic growth model for species group 3.

This resulted in changes in the number of stems and total basal area for a range of tree sizes and time intervals. This shows that the tendency for individual trees to grow either faster or slower than the mean predicted by the deterministic model is significant. The observed changes can be attributed to creating an imbalance between growth, recruitment and mortality and the time that any one tree is expected to remain in one of the standard size classes.

Test parameter set ten altered the spread of the distribution of the stochastic variation in predicted growth in the model. The results showed that as this spread increased, the sum of basal area tended to increase after 50 years of simulation. This was a result of the cumulative effect of enhanced growth on large trees.

Test parameter sets 12, 14 and 15 demonstrate that predictions from the ecological model are sensitive to the parameter values used in the mortality model. These are the most significant

Table 5
The results showing significant systematic effects

Numbers	<i>j</i>	Size class	Species	Time	n_c	$n: F_{i,j}$	$v: F_{i,j}$	$b: F_{i,j}$
1	8	1	3	35	26.8	0.4		−4.7
2	8	1	3	40	25.7	−0.1		−5.9
3	8	2	3	40	13.1	−4.0		−4.5
4	8	2	3	50	14.2	−4.8		−4.6
5	8	1	3	75	19.0	4.6		−1.7
6	8	2	3	75	14.0	−5.6		−6.0
7	8	1	3	100	18.2	6.0		1.0
8	8	2	3	100	11.4	−4.0		−3.8
9	8	3	3	100	8.0	−5.6	−2.9	−5.2
10	10	1	8	50	244.6	1.9		4.1
11	10	2	8	100	42.4	4.3		4.2
12	12	5	2	30	28.9	−2.0	−2.8	−4.3
13	12	5	2	75	28.8	−3.5	−2.8	−4.3
14	12	5	2	100	28.5	−2.3	−2.9	−4.5
15	14	5	2	30	28.9	1.6	4.0	5.9
16	14	5	2	35	29.3	1.3	4.0	5.6
17	14	5	2	40	29.1	1.8	4.7	6.7
18	14	5	2	50	28.9	0.5	4.5	6.1
19	14	5	2	75	28.8	0.5	3.2	4.7
20	14	5	2	100	28.5	0.6	3.3	5.7
21	15	5	3	20	46.3	0.2	0.2	4.3
22	20	1	2	100	14.0	4.9		3.7
23	20	5	2	100	28.5	4.7	−0.1	0.4
24	21	1	3	20	29.3	−4.6		−3.4
25	21	1	3	30	28.0	−4.3		−3.2
26	21	5	3	30	47.3	−4.2	−0.7	0.5
27	21	1	3	40	25.7	−4.3		−3.3
28	21	5	3	40	47.1	−4.2	−0.4	−0.2
29	21	1	3	50	23.1	−4.9		−2.7
30	21	5	3	50	46.3	−5.4	−0.4	−0.6
31	21	1	3	75	19.0	−8.3		−6.7
32	21	5	3	75	44.1	−8.4	0.3	−1.8
33	21	1	3	100	18.2	−9.4		−7.7
34	21	5	3	100	42.9	−11.6	−0.5	−2.5

j Indicates the test parameter set that gave the result, n_c is the number of trees in the control result and $F_{i,j}$ are the differences between the control and test results as a proportion of the statistical error, for variables n , the number of trees, v , the sum of the stem volume for commercial trees and b , the sum of the basal area. The blank entries for volume are for small size-classes for which there can be no commercial stems.

results from this analysis as the results for total basal area and commercial volume were affected, and it is these values that are of most interest to users.

The significant results obtained for test parameter sets 20 and 21 show that the number of trees in the smallest size class is sensitive to the parameters in the recruitment model, but that these effects do not tend to be observed in either

the total basal area or commercial volume in simulations of up to 100 year duration. It would be expected that these effects might become significant during longer simulations.

The remaining results from this analysis were examined to detect more subtle effects. For the case of no systematic effects, the mean should be consistent with zero and the standard deviation (the spread of results) should equal the statistical

error, which has been normalised to 1.0. For the test sets the mean was -0.039 ± 0.018 and the standard deviation was 1.0721 ± 0.018 , demonstrating a small systematic effect. Using Eq. (22), the systematic error associated with the increase in spread was calculated to be 0.38 times the size of the stochastic error from 30 repetitions. Thus the random or stochastic error dominated the total error for most results. This outcome emphasises the need for users to implement sufficient repeat simulations in any study.

The results from the analysis of systematic errors were combined and presented in a form that could be used in evaluation of simulation results to estimate systematic errors for specific results. This information is available in the help files of the model and the SYMFOR web site.

4.3.2. General guidelines relating to systematic errors in the ecological model

The analysis of systematic errors for the ecological model demonstrated the following issues:

- Systematic errors are relatively insignificant for the first 30–50 years of any simulation. Their effect increases with time and may become much more significant at the end of a 100 year or longer simulation.
- Results from the model are most sensitive to the values used in the mortality and recruitment sub-models. This result reflects the relative lack of data available to develop and calibrate these components of the model.
- The deterministic growth model parameters did not give significant systematic effects. There were more data to calibrate this model than the mortality or recruitment models, so resulting uncertainties were smaller.
- Systematic errors in the estimation of mortality will strongly influence basal area and volume estimates of relevance to forest managers. An underestimate of mortality will result in increased total basal area and volume. This is particularly important for the largest trees.
- Systematic effects tend to be larger for the faster-growing species (group 3) and canopy species (groups 2 and 3) than they are for the slower-growing sub-canopy species (group 8),

even though the sub-canopy species dominate in terms of the number of trees. This is a result of the representation of stand dynamics in the model.

The most effective way of reducing the effect of systematic errors on results from simulations is to calculate estimates using paired silvicultural trials, for example a conventional and alternative management regime. This approach then allows statistical analysis of the differences between pairs of estimates using an analysis equivalent to a paired *t*-test. The sources of systematic error are the same for both trials, and it would be expected that the majority of the effects would be the same for both treatments. This approach has the added advantage of being able to remove the effects of variation between plots used to initialise each simulation, and is the recommended method of application of the SYMFOR framework to support management decisions.

5. Range of validity for the model

Descriptive (statistical or regression) models that have been derived from data should, in general, not be used to make predictions outside the range of conditions that were used during the process of model calibration. Ideally it would be possible to classify the system by one or more numerical quantities in order to determine whether a particular application of the model is within the range of validity. For complex biological systems

Table 6
The numerical validity indicators and the range of their values covered by the data used for calibration

Validity indicator	Minimum value	Maximum value
Diameter of trees (cm)	10	171.7
Shade index at 5 m	0.0	25.5
Shade index at 30 m	0.75	42.9
Years since logging	1	Unlogged
Shade index at 30 m averaged over a grid-square	3.2	36.4

such as models of tropical forests, this process may be to some extent subjective.

The range of numerical conditions encompassed by the data used to develop the model is shown in Table 6. These values are of most relevance to the growth model, but do not describe the impact of environmental factors such as soil type, topography and climate. These can only be evaluated qualitatively, and are described in Section 3.1 for the data used. The species composition of the forest being simulated is extremely important and should be compared with published reports for the STREK plots (Bertault and Kadir, 1998) and others in the region (van Gardingen et al., 1998). Ideally the model should be tested against data from other regions before producing results, and where this is not possible, a qualitative statement should be made regarding the model validity.

6. Performance of the model compared against observed data

As stated in the introduction, the purpose of this model was to be initialised with real data and then used to predict stand structure and composition for a period of between 10 and 100 years. The model is required to make predictions about the growth of individual trees and aggregated characteristics describing the stand. The performance of the model can be described through the analysis of individual model components (growth, recruitment and mortality) and the change in the characteristics of the stand with time. Both approaches are described here.

6.1. Performance of individual model components

The performance of each model component was compared against the data observed in the STREK plots. These comparisons are combined here with a general discussion of the model.

6.1.1. The deterministic growth equations

The performance of the deterministic growth model was assessed by the proportion of variation in the dataset explained by the model as estimated by the coefficient of determination (R^2). The R^2

value for the whole model, including the species grouping and the deterministic model equation, was 35.6%. The proportion of this variation that could be explained by the species grouping alone (in the absence of the growth model) was calculated to be 13.6%. The growth models were applied separately to each of the species groups to give the values shown as Table 8, with a maximum of 74.9%, and a minimum of 12.3%, with a mean of 25.6%. This table also shows the size of each species group specified as the percentage basal area and number of stems compared with the whole data set. This shows that R^2 tends to be lowest for the largest groups.

The values of R^2 calculated above compare favourably with Rombouts (1998) descriptions of the development of five alternative groupings of the species and growth equations for individual trees in the DIPSIM model, using the same data as in this study. Gourlet-Fleury and Houllier (2000) evaluate models for each of their 15 species groups for forest in French Guiana. They calculate R^2 only from the 87 species used to produce the species groups, whereas this study has used data from all trees, so the figures are not comparable (although the values obtained are similar). Other studies (Kleine and Heuvelodp, 1993; Ong and Kleine, 1996; Huth et al., 1997, 1998; Köhler and Huth, 1998) do not report R^2 values for their models.

Overall the model explains over 35% of the variation in growth rates observed in the STREK dataset. The remaining variation is likely to have resulted from genetic variation between individuals and between species within a group, measurement error, and the main environmental factors known to influence growth but not recorded in the dataset and hence not represented in the model. Better performance would not necessarily be expected, because these known causes of variation in growth could not be modelled because suitable data were not available.

6.1.2. The stochastic growth component

The deterministic growth model estimates the correct average growth rate for given values of species group, size, tree position and local silvicultural history. The difference between predicted

values and observed data form a distribution around this value. This distribution of growth rates is ecologically significant. For example, in a dense monoculture of equally spaced trees of uniform size, the deterministic model would grow all trees equally. In practice, this is not the case: genetic variation, local soil characteristics, specific crown topology and other stochastic events will lead to some trees growing faster than others, obtaining an initial advantage and consequently dominating the forest. These effects represent competition between individuals and are responsible for variation in growth rates and the resulting sizes of individuals even in so-called uniform forest stands. An ecological forest model should be able to simulate these effects. In this simple example, modelling the remaining variation in growth rate is essential to determine the future structure of the forest because differences between trees become compounded through competition over time. The stochastic part of the model describes the remaining variation, but does not attempt to explain it. The importance of the stochastic component of the growth model was demonstrated in the results of the estimation of systematic errors (Section 4).

The concept of serial correlation (Vanclay, 1994) is intrinsically linked to the idea of differences between trees' growth rates. The notion of the stochastic distribution being temporally consistent enhances the compounding effect. Trees that happen to be in favourable locations (with respect to soil characteristics, etc.) should consistently grow faster than ones that are not. Significant serial correlations were found for all species groups, as shown in Table 1 (parameter p_{10}). This demonstrates that ignoring serial correlation would be to homogenise growth rates towards the deterministic model average.

A limitation in developing the current model was that one might expect trees with consistently higher (or lower) growth than the average to be clustered in particular locations, such as ridge-tops. If suitable data were available that described such environmental variation, these effects could be included in the deterministic model. The presence or absence of these locations is not known from the data, and thus the variation was modelled as part of the stochastic variation.

6.1.3. The recruitment model

Recruitment into the modelled plot is stochastic, representing the fact that there are no data describing trees with DBH less than 10 cm: the information necessary for deterministic modelling.

Grid-squares with a side length of 10 m were chosen in order to represent recruitment across the plot, whilst keeping recruitment probabilities per species group, per grid-square, per year, making a compromise between computation time and excessive heterogeneity of the hypothetical seedling distribution. The process of tree recruitment into a size category above 10 cm depends not only on the conditions at the current time, but also those in the past, for example previous soil disturbance and the presence of seed. The model works with the assumption that soil disturbance removes all existing seedlings and that there is a supply of seed (or a seed bank) for all species groups.

A further factor affecting the final values of the recruitment parameters was that species group 10 is assigned to trees of unknown taxa for purposes of growth and mortality but not for recruitment. Group 10 is, by definition, a mix of species from other groups, and it may be assumed that these trees could have been assigned to another species group that more precisely matched the behavioural characteristics of that species. Hence there is no reason to simulate recruitment of trees into group 10. Instead, the parameter values for recruitment for species groups 1–9 were each scaled by the proportion of new trees of group 10 to the total number of new trees observed in the data. Thus, for a forest whose development has been simulated for many tree generations one would not find any trees of species group 10.

6.1.4. The natural mortality model

The processes of natural mortality resulting from causes such as windfall, drought, disease or soil destabilisation were modelled stochastically. For an explicit, spatial model of forest dynamics, however, it is also important to model gap creation and the damage resulting to other trees in the stand. When a large tree falls, it will damage or kill smaller trees in its path. The damage is, to some extent, predictable, and mortality from tree-fall damage is modelled separately.

The section of calibration of the model noted that insufficient data existed to obtain adequate calibration of the mortality model. An alternative approach was adopted where the parameters for the mortality model were adjusted to generate a dynamic equilibrium with growth and recruitment for primary forests. It is obviously preferable to have sufficient data to calibrate and test the mortality model, and the importance of this is emphasised by the results from the estimation of systematic errors in Section 4. Natural mortality in forests is a relatively rare event, and hence the only way of obtaining better data is through the long-term monitoring of permanent sample plots.

6.2. Changes in stand structure over time

The performance of the model can be evaluated against time series obtained from permanent sample plots. The available data from the STREK plots covered a period of 6 years in three areas of forest. One described 12 ha of primary forest that remained unlogged over the entire measurement period. Another area of primary forest (12 ha) was logged between the first and second measurements, giving a period of 4 years regrowth immediately following logging. The third area (8 ha), which had been logged in 1980 and had not been treated since, represented forest that had a longer period of recovery following logging. The time series from these plots was used to test the performance of the model.

6.2.1. Method

Simulations were run for each set of plots. The simulations ran for 6 years for the primary forest plots and those first logged in 1980. The 12 ha plots that were logged in 1992 were used to represent recently logged forest and were simulated for the 4 year period immediately following the management intervention up to the final measurement of the plots.

The performance of the model was evaluated using the same set of standard result variables described for the evaluation of systematic errors (Section 4) after the size-class representing all trees was excluded. The difference of these quantities between the initial (at the first measurement) and

final value (at the last measurement or end of the simulation) were then averaged over all plots of similar silvicultural history to obtain the mean and standard error of the change in value of each variable over the period. Ten repeat simulations were run for each plot and the mean result of these used to calculate the mean change over all plots. The difference between the mean data value and the mean predicted value was calculated and divided by the combined error term, in order to examine the distribution of results for each result variable. This produced a standardised difference providing a frequency distribution of the results in terms of the expected deviation from zero that could be compared with a Gaussian distribution.

6.2.2. Results

Table 7 shows the percentage of results that are found within \pm a given number of standard deviations of zero. It should be noted that this is equivalent to a cumulative normal distribution. Normal distribution expected 95% of observations to be less than ± 2 standard deviations from the mean. The results for the commercial volume were close to this value and were achieved for basal area results. The main deviation from the expected results was observed for the number of trees in each size class. The earlier analysis of systematic errors had also identified that these results were likely to be influenced by the recruitment and mortality models.

The result sets which demonstrated large deviations (greater than 2 standard deviations) were examined. These were found to mainly occur for

Table 7

The percentage of results within a number of standard deviations (N_σ) of the mean, for number of trees (N_N), basal area (N_B) and volume (N_V) compared with that expected for a perfect normal distribution (N_E)

N_σ	0–1	1–2	2–3	3–4	4–5	5–6	6–7
N_E	68	95	99	100	100	100	100
N_N	58	78	89	91	92	99	100
N_B	57	95	97	99	100	100	100
N_V	62	93	99	100	100	100	100

Table 8

The R^2 values for the deterministic model of growth within each species group, with the percentage of the total basal area, B , and percentage of the total number of trees, n , represented by each species group shown to indicate the relative importance of the groups. It should be noted that the species groups themselves explain 13.6% of the variation observed in the whole data set. The data used for this were the same as used in model development

Species group	R^2	B	N
1	40.6	2.5	1.6
2	12.3	12.4	5.0
3	19.3	22.4	8.9
4	18.1	7.8	8.2
5	74.9	0.1	0.1
6	33.5	1.4	2.2
7	13.2	9.9	13.0
8	15.9	29.1	38.5
9	14.6	5.5	8.4
10	15.2	8.9	14.1

the area of forest that had been logged 11 years before the first data were collected, and entirely in the smallest size-class of trees. They occur for species groups 2, 4, 5, 6, 7, 8 and 9, and are in all cases a prediction of larger increases in numbers of trees than were observed in the data. This again illustrates that the recruitment model has a large effect on the total number of small trees. The smallest size-class dominates the total number of trees of a species, and is primarily influenced by the processes of recruitment and mortality rather than growth. The earlier section showed, however, that this is unlikely to effect results of significance to forest managers over the period of most simulations. The simulation results for larger size-classes were in good agreement with the observed data, and it is these size-classes that are important to forest structure and dynamics.

7. Application of the model

The final discussion of the model is illustrated here through sample simulations. Other aspects of the application are discussed in more rigorous applications published elsewhere as technical notes

(<http://www.symfor.org/technical>) and research papers (van Gardingen et al., 2002).

7.1. Simulations

A simulation of forest behaviour following heavy disturbance was made to demonstrate the combined effects of the growth, recruitment and mortality modules. A plot of forest logged in 1992 was used. The original logging had reduced the stand from 427 trees (with a DBH of more than 10 cm) to 247. Data collected after logging were used to initialise the model which was implemented with all management (harvesting) options disabled. Data were output to enable calculation of two quantities: the total basal area and the number of trees, for each species group for each year of the simulation.

Figs. 2 and 3 show the variation of the number of trees and the sum of the basal area with time for each species group. The model simulates significant increases in the number of trees through recruitment. The largest numbers of new recruits were representative of the pioneer species in groups 5 and 6. The number of pioneers began to decline after approximately 15 years as would be expected with canopy closure, but the sum of

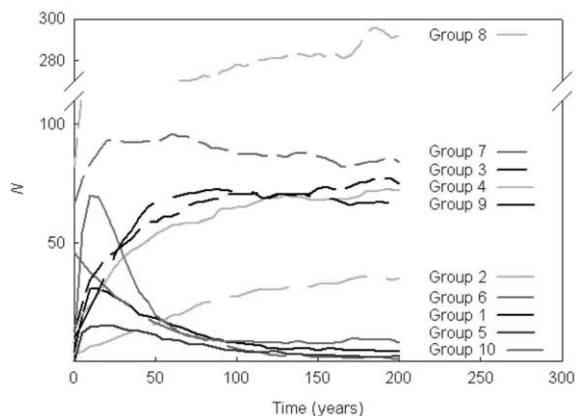


Fig. 2. The number of trees, N , of each species group as a function of time since the simulation began for a heavily disturbed plot. These results are the average of ten replicate runs, to reduce stochastic variation.

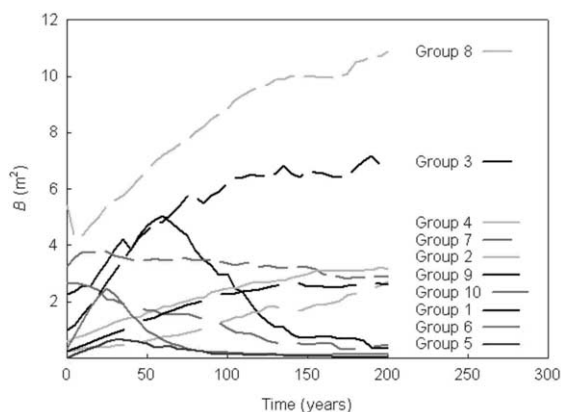


Fig. 3. The sum of the basal area, B , for each species group as a function of the time since the simulation began, for a heavily disturbed plot. These results are the average of ten replicate runs, to reduce stochastic variation.

the basal area continued to rise to a peak at around 25 years, reflecting the growth of surviving individuals. Group 8, representing small understorey trees, dominates the forest in terms of both numbers of trees and total basal area. It should, however, be noted that this group is smaller than the sum of all Dipterocarp species (groups 1–4) in the forest. The results also show that fast-growing *Shorea* species (species group 1) exhibits similar recruitment behaviour to the pioneers with rapid increases in numbers soon after logging. These species are known to be more light-demanding than other Dipterocarp species and would become established in the disturbed areas created during logging. The total basal area of the fast growing *Shorea* group continued to increase to around 60 years reflecting the greater longevity and size of trees in this group. The remaining Dipterocarp species (groups 2–4) tend to regenerate under relatively closed canopy and demonstrate a continued recovery of both numbers and basal area after the initial effects of disturbance.

These results predict very significant changes in the species composition of the forest in terms of both numbers of individuals and their basal area over the 200 years of simulation. The initial shift is towards more rapidly growing light-demanding species that become established in disturbed areas.

The species are gradually replaced by shade-tolerant species as the simulation progresses and canopy closure changes the environment for recruitment. The results show that the forest could be expected to still be in a state of change after 200 years, stressing the long time scale of likely responses to disturbance.

The data used to calibrate the model included a significant number of individuals that had not been identified; these were placed in species group 10. It was noted previously that the recruitment model excluded this group, by assigning new trees to one of the other nine groups. The total number of trees in this group and their basal area declines with time for this reason. One consequence of this is that the number of individuals in all other groups should be expected to increase in time, even in unlogged forest.

It is noted that in typical unlogged forest there are more individuals of species 1, 5, 6 and 7 than are observed in the simulated forest after 200 years. This can be attributed to the lack of representation of catastrophic natural mortality in the model. Catastrophic mortality includes damage from storms, droughts, landslides, earthquakes and other infrequent large-scale disturbances that would cause more widespread damage at a point in time than normal natural mortality. Such disturbances would create suitable conditions for the growth of a few individuals of species groups 1, 5, 6 and 7, as observed in real plot data. It would also cause the removal of large individuals of the climax species groups. The long-term effect of these disturbances would be to curtail the apparent convergence of the simulated forest towards a steady state.

7.2. Variability between simulations and plots

The model is a stochastic model and thus requires multiple simulations to derive mean estimates of the expected response for any one plot. The estimation of the random error associated with the stochastic components of the model was discussed with the analysis of systematic errors. It was shown that simulations should be repeated for each plot and that this may require 10–30 simulations to be completed per plot.

Analysis of the data used to develop the model also demonstrated very significant variation between plots with the average basal area ranging from 27 to over 37 m² ha⁻¹ for different areas of primary forest at the first survey. The calibration process produces models that tend to represent the behaviour of an average plot rather than any particular plot. This means that characteristics of plots will tend to converge on average values as the simulation progresses. It is expected that the variation between plots will decrease, especially for long simulations. The combined effect of these characteristics of the application of the model (variation between plots and the model converging on average values) reinforces the previous statement that the most reliable way to use the model is to make paired simulations of treatments for each plot or set of plots.

7.3. Representing disturbance in the model

Disturbance over varying spatial and temporal scales is responsible for dynamic equilibrium of the tropical forests and their structure. The extent of disturbance ranges from relatively frequent creation of small gaps from individual tree falls through to infrequent catastrophic events that may affect many hectares of forest. The dynamics of the forest also respond to extreme climatic events such as drought and storms, which may have average return periods of many years or decades.

The ecological model was developed with data collected over a relatively short time period and for this reason the data do not describe the effects of any extreme disturbance events. The model at present only simulates the effects of frequent small-scale disturbance such as may result from either natural tree death, or from normal harvesting operations. Additional data from a longer series of measurements would be required before any other type of disturbance could be simulated.

Data collection for the frequency and intensity of the catastrophic mortality events would require observation over a period of decades or centuries, and is not available for the Berau dataset used to calibrate this model. The island of Borneo is known to experience severe droughts at intervals

in conjunction with *El Niño* events. Such a regional event was observed during the period 1997–1999. Data describing the effects of the drought are now available for the STREK plots. Statistical analysis has shown that there were reduced growth rates and enhanced mortality during this period. It has not yet been possible to include these effects in the ecological model as there is no information to estimate the severity of the drought in comparison with previous events, nor is there any information that could be used to estimate the frequency of these events. The growth, recruitment and mortality functions in the ecological model could be modified to account for the effects of extreme events. Alternative data sets from the island of Borneo are being investigated for this purpose. In the interim period, the effects of drought and other extreme events will be included through the long-term calibration of model components.

7.4. Captured ecology

The expert knowledge of field ecologists presents one of the most important qualitative assessments of an ecological model. The model should make predictions that are consistent with field observations. Many of the components in the model were developed from expert knowledge, for example the delay in recruitment along compacted skid trails. Other aspects result from complex interactions such as the disturbance regime following logging and subsequent recruitment and growth in the disturbed areas. The model predicts the behaviour of individual trees whilst the behaviour of the stand as a whole becomes an emergent property of the model. The performance of the model can then be evaluated by comparing the behaviour of the stand with predictions made by forest managers and ecologists.

The results from the model shown as Figs. 3 and 4 show the expected behaviour of the forest following heavy logging. Several species groups increase in number of trees and total basal area immediately following logging. The fast growing *Shorea* species (group 1) and pioneer species (groups 5 and 6) become important. These effects are spatial and reflect the ecology of gap-phase dynamics (Brokaw, 1985; Denslow, 1987) in forest

succession. For example, the most light-demanding species, *Anthocephalus chinensis* (Group 5) only becomes established in highly disturbed areas in the model, such as a region where several skid trails converge. This characteristic of modelled stand behaviour is identical to that observed in the field. *A. chinensis* is used as a field indicator of excessive disturbance following logging. The ecological model simulates this well, as can be shown by simulating good management using the combination of yield regulation and reduced impact logging. The modelled behaviour of the forest matches field observation in that very few (if any) individuals of *Anthocephalus* become established.

The dynamics of a forest are dominated by few large individuals. The death of one large tree will have significant impact on the total basal area of that species, whilst the tree falling will damage or kill a large number of smaller trees. The design of the model utilises spatial information for individual trees and represents these processes well. The importance of the few large trees is one reason that so many repeated runs are required to obtain reliable statistical estimates from the model. Exactly the same issues are prevalent in field studies.

There are, however, many aspects of ecology currently not represented by the model. These include the environmental effects of soil type, topography, aspect and biological effects such as mycorrhizal associations or genetic systems. In each case, these processes could be incorporated into the model if suitable data were available for model development and calibration. Any such work should, however, consider that the analysis of systematic errors in this study identified the mortality component of the model to be associated with the greatest uncertainty of relevance to the intended users of the model, with recruitment processes becoming important for long-term simulations.

8. Conclusions

The model described in this paper meets the requirements established in the introduction; it is based on statistical analysis of complex time series data from Dipterocarp forest in Indonesia and

presents this information in a form that can be used to inform decisions by forest managers and policy makers. The model also meets the requirement of capturing much of the ecological characteristics of the forest so that those aspects can be used to inform decisions.

The model described here differs in two significant ways from most previous studies. Firstly, the model is based on the description of the growth of individual trees in a spatially explicit environment. Secondly, the ecological model is incorporated into the SYMFOR framework so that it can interact with models of management treatments. This approach permits an analysis of the ecological consequences of logging to be combined with predictions of future growth and yield. One example given was the changes in the proportion of pioneer species along skid trails following logging.

The framework is now fully integrated with ecological and management models and is being used to investigate the effects of alternative silvicultural practices on the forest into the future. This tool can provide insights to the likely outcomes from silvicultural practice that are not possible to achieve with experimentation, due to the timescales involved. It relies completely on the ecological model for this purpose. The tests and validations discussed here provide the users with estimations of the reliability and accuracy of their results and outline the range of conditions in which the model and its predictions are likely to be valid. The design of the framework is such that this approach can be adapted to other areas by modifying the ecological model. The simplest method is to calibrate the model using new data from permanent sample plots, with an alternative being available to replace one or more components of the ecological model with new modules derived from new data.

The simulation presented in this paper demonstrated the breadth of ecological and management information that can be obtained from the application of the model. These are as follows:

- The process of calibrating the model requires an integrated statistical analysis of growth, recruitment and mortality data, which makes this

information available in a way that is often much easier to interpret than conventional statistical analysis. It allows objective comparisons of ecological characteristics between species groups or location and forest types.

- Interpretation of results from simulations can reveal information describing the ecology of recovery after disturbance such as harvesting. The simulation presented in this paper demonstrated that there is a rapid increase in the number of individuals of pioneer species and their total basal area after logging. The strength of the ecological model is that it allows the prediction of ecological processes in the forest resulting from management interventions. Other simulations (not presented) have shown that frequent logging leads to a change in species composition with more fast-growing dipterocarps and pioneer species.
- The model can be further used to present aspects of forest ecology to a range of stakeholders. One such application is to use the model and the SYMFOR framework to enhance the ecological component of forest curricula in universities and training organisations. An ecological model is the only effective way to demonstrate the interactive effects of individual processes of growth, recruitment and mortality.
- The ecological model allows users to examine the likely impact of forest management interventions on ecological aspects of the forest. This is a key process required to support the development of more sustainable forest management systems for tropical forests.

The analysis of the performance of the model and the systematic errors associated with simulations gives guidelines for the application of the current model and areas for future improvements. The application of the model for management purposes over a period of less than 50 years is likely to produce accurate predictions of growth, yield and species composition. Simulations over longer periods are likely to be less accurate because of uncertainties associated with the mortality and recruitment components of the model. The impact of these uncertainties are minimised when the model is used to quantify the differences

between treatments through paired simulations using the same sets of plots. Any potential error will be further minimised when the disturbance resulting from logging is not extreme such that it is similar to that found in natural forests (single tree falls). These observations mean that the model can be used with confidence to inform management decisions. Estimates of yield should be accurate for at least 50 years, whilst simulations of longer periods will generate realistic trends.

Future developments of the model will be entirely dependant on the availability of additional data. The existing components of the model can be recalibrated when more data become available from the STREK plots or suitable data become available from other regions. The mortality and recruitment model components have been identified as having the highest priority for new work. The nature of these components is such that suitable data will only become available from long-term datasets. The generic form of the model means that it can be applied to other locations and forest types as long as suitable data are available for calibration. This process represents the next significant challenge for this research and is currently being addressed for forests in South America and Africa.

Acknowledgements

The authors would like to acknowledge the staff of P.T. Inhutani 1, BPK Samarinda and the European Union Berau Forest Management Project for the provision of the data necessary for this work, assistance in data processing and their comments on this manuscript. This publication is an output from a project funded through the Forestry Research Programme of the United Kingdom Department For International development (DFID) for the benefit of developing countries. The views expressed are not necessarily those of DFID.

Appendix A

The meaning, units (where appropriate) and range of possible values for each variable used in the model description. The variables are introduced within the text, but are collated here for ease of reference.

Variable	Meaning/context	Units	Value range
H	tree height	m	> 0.0
s	coefficient in height equation	–	$\neq 0.0$
D	tree diameter at breast height (1.3 m)	cm	> 10
H_m	coefficient in height equation	m	$\neq 0.0$
C_p	height of mid-point of tree crown	m	> 0.0
f_c	coefficient in crown-point equation	–	> 0.0
C_R	radius of tree crown at mid-point	m	> 0.0
f_r	coefficient in crown-radius equation	–	> 0.0
B	stem basal-area	m ²	> 0.0
ΔD_p	annual diameter increment (deterministic)	cm	≥ 0.0
ΔD_S	annual diameter increment (stochastic)	cm	≥ 0.0
S	shade index at 5 m	m ⁻¹	≥ 0.0
D	distance between two trees	m	> 0.0
S'	shade index at 30 m	m ⁻¹	≥ 0.0
G_d	diameter term in growth equation	cm	any
G_S	shade index at 5 m term in growth equation	–	any
$G_{S'}$	shade index at 30 m term in growth equation	–	any
G_Y	years since logging term in growth equation	cm	any
p_0	coefficient in G_d equation	cm	any
p_1	coefficient in G_d equation	–	any
p_2	coefficient in G_d equation	cm ⁻²	≥ 0.0
p_3	coefficient in G_d equation	cm	any
p_4	coefficient in G_S equation	m ²	≥ 0.0
p_5	coefficient in G_S equation	m	any
p_6	coefficient in G_Y equation	cm per year	any
p_7	coefficient in G_Y equation	year	≥ 0.0
F_Y	scaling term in G_Y equation	–	≥ 0.0
N_T	number of live trees following logging	–	≥ 0
p_8	average number of trees before logging	–	≥ 0
r	residual of actual growth and ΔD_d of a tree	cm	any
ΔD_O	observed annual growth of a tree	cm	≥ 0.0
g_r	a possible, modelled value of r .	cm	any
$P(g_r)$	probability of a particular g_r occurring	–	0.0–1.0
p_{11}	coefficient in $P(g_r)$ equation	cm	≥ 0.0
p_{12}	coefficient in $P(g_r)$ equation	cm	any
R^2	a statistical measure of model quality	–	0.0–1.0
r_{1-6}	random numbers	–	0.0–1.0
$N_{a,b}$	percentage of similar pairs of ΔD_O observed	–	%
N_e	the random distribution value of $N_{a,b}$	–	%

p_{10}	serial correlation between residuals, r	–	0.0–1.0
$Q(g_r)$	hill-top probability distribution of g_r	–	≥ 0.0
g_1	the lesser of p_{10} and $1-p_{10}$	–	0.0–0.5
Q_m	minimum value of $Q(g_r)$	–	≥ 0.0
g_m	minimum possible g_r	cm	any
$P'(g_r)$	corrected form of $P(g_r)$	–	0.0–1.0
P_i	annual recruitment probability for species i	–	0.0–1.0
m_i	coefficient in recruitment equation	cm^{-1}	any
c_i	coefficient in recruitment equation	–	any
t_i	the in-growth time for species i	year	> 0
P	annual mortality probability for a tree	–	0.0 – 1.0
m_1	coefficient in mortality equation	–	any
m_2	coefficient in mortality equation	cm^{-1}	any
D_m	threshold diameter in mortality equation	cm	any
m_3	probability of a dying tree falling over	–	0.0–1.0
α_f	fall-direction of fallen trees	rads	0.0– 2π

Appendix B The characteristics and content summary of each species group.

Group	Name (reference)	Characteristics	Dominant members
1	fast growing <i>Shorea</i>	Large trees, light demanding, very fast growing	<i>Shorea johorensis</i> , <i>S. leprosula</i>
2	<i>Dipterocarpus</i>	Large trees, shade tolerant, slow growing	<i>Dipterocarpus</i> , some <i>Shorea</i>
3	other large dipterocarps	Large trees, shade tolerant, fast growing	<i>Shorea</i> , <i>Parashorea</i> , <i>Dryobalanops</i>
4	small <i>dipterocarps</i>	default group for dipterocarpaceae species	<i>Hopea</i> , <i>Vatica</i> , <i>Dipterocarpus</i> some <i>Shorea</i>
5	<i>Anthocephalus</i>	small trees, fast growing, highly disturbed forest	<i>Anthocephalus chinensis</i>
6	<i>Macaranga</i>	small trees, light demanding, very fast growing	<i>Macaranga</i>
7	gap small trees	small trees, recruit in light areas	<i>Aglaia</i> , <i>Knema</i> , <i>Artocarpus</i>
8	other small trees	small trees, default group for non-Dipterocarpaceae species	<i>Diospyros</i> , <i>Dacryodes</i> , <i>Polyalthia</i>
9	shade small trees	small trees, recruit in shady areas	<i>Macaranga lowii</i> , <i>Gonystylus</i> , <i>Madhuca</i> , <i>Kayea</i>
10	unknown	‘unknown’ species, genus or family identity	unknown

References

- Alder, D., Synott, T.J., 1992. Permanent sample plot techniques for mixed tropical forests. Tropical Forestry paper 25, Oxford Forestry Institute, Department of Plant Sciences, University of Oxford, UK.
- Barlow, R.J., 1989. Statistics: A Guide to the Use of Statistical Methods in the Physical Sciences. Wiley, Chichester, UK.
- Bertault, J.G., Kadir, K., 1998. Silvicultural Research in a Lowland Mixed Dipterocarp Forest of East Kalimantan. Cirad, Montpellier, France.
- Bertault, J.G., Sist, P., 1997. An experimental comparison of different harvesting intensities with reduced-impact and conventional logging in East Kalimantan, Indonesia. For. Ecol. Manage. 94, 209–218.
- Brokaw, N.V.L., 1985. Gap phase regeneration in a tropical forest. Ecology 66, 682–687.
- Clark, D.A., Clark, D.B., 1992. Life history diversity of canopy and emergent trees in a neotropical rainforest. Ecol. Monogr. 62, 315–344.
- Denslow, J.S., 1987. Tropical rain forest gaps and tree diversity. Annu. Rev. Ecol. Syst. 18, 431–451.
- Departemen Kehutanan R.I., 1993. Pedoman dan petunjuk teknis Tebang Pilih Tanam Indonesia (TPTI) pada hutan alam daratan. Departemen Kehutanan, Direktorat Jenderal Pengusahaan Hutan, Jakarta, Indonesia.
- Dykstra, D.P., Heinrich, R., 1996. FAO Model Code of Forest Harvesting Practice. Food and Agriculture Organisation of the United Nations, Rome, Italy.
- Elias, 1998. Reduced Impact Timber Harvesting in the Tropical Natural Forest in Indonesia. Forest Harvesting Case Study 11. Food and Agriculture Organisation of the United Nations, Rome, Italy.
- Fishman, G.S., 1995. Monte Carlo: Concepts, Algorithms and Applications. Springer, London.
- Gourlet-Fleury, S., Houllier, F., 2000. Modelling diameter increment in a lowland evergreen rain forest in French Guiana. For. Ecol. Manage. 131, 269–289.
- Huth, A., Ditzer, T., Bossel, H., 1997. Rainforest growth model FORMIX3: a tool for forest management planning towards sustainability. Model development and case study for Deramakot Forest Reserve in Sabah, Malaysia. TOB-series number: TOB FTWF-6/e. Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) GmbH, Eschborn, 78 S.
- Huth, A., Ditzer, T., Bossel, H., 1998. The rainforest growth model FORMIX3—model description and analysis of forest growth and logging scenarios for the Deramakot Forest Reserve (Malaysia). Gottinger Beitrager zur land- und Forstwirtschaft in den tropen und subtropen 124. Erich Goltze, Gottingen, 182 S.
- Kirschbaum, M.U.F., 2000. Forest growth and species distribution in a changing climate. Tree Phys. 20, 309–322.
- Kleine, M., Heuvel, J., 1993. A management planning concept for sustained yield of tropical forests in Sabah, Malaysia. For. Ecol. Manage. 61, 277–297.
- Köhler, P., Huth, A., 1998. The effects of tree species grouping in tropical rainforest modelling: Simulations with the individual-based model FORMIND. Ecol. Model. 109, 301–321.
- Köhler, P., Ditzer, T., Ong, R.C., Huth, A., 2001. Comparison of measured and modelled growth on permanent plots in Sabahs rain forests. For. Ecol. Manage. 144, 101–111.
- MacFarlane, D.W., Green, E.J., Valentine, H.T., 2000. Incorporating uncertainty into the parameters of a forest process model. Ecol. Model. 134, 27–40.
- Newbery, D.M., Kennedy, D.N., Petol, G.H., Madani, L., Ridsdale, C.E., 1999. Primary forest dynamics in lowland dipterocarp forest at Danum Valley, Sabah, Malaysia, and the role of the understorey. Philosophical Transactions of the Royal Society of London Series B. Biol. Sci. 354, 1763–1782.
- Ong, R., Kleine, M., 1996. DIPSIM: a dipterocarp forest growth simulation model for Sabah. FRC research papers No. 2. Forestry Department, Sabah, Malaysia.
- Phillips, P.D., van Gardingen, P.R., 2001. The SYMFOR framework for individual-based spatial ecological and silvicultural forest models. SYMFOR Technical Notes Series No. 8, <http://www.symfor.org/technical/index.html>, The University of Edinburgh, UK.
- Phillips, P.D., Yasman, I., Brash, T.E., van Gardingen, P.R., 2002. Grouping tree species for analysis of forest data in Kalimantan (Indonesian Borneo). For. Ecol. Manage., in press.
- Rombouts, J., 1997. Data structure of the growth and yield clearing house. Department for International Development, Jakarta, pp. 1–20.
- Rombouts, J., 1998. Species Grouping Based on Diameter Increment in East Kalimantan. SFMP, Samarinda, Indonesia.
- Saltelli, A., Tarantola, S., Chan, K.P.S., 1999. A quantitative model-independent method for global sensitivity analysis of model output. Technometrics 41, 39–56.
- Sist, P., Bertault, J.G., 1998. Reduced impact logging experiments: impact of harvesting intensities and logging techniques on stand damage. In: Bertault, J.G., Kadir, K. (Eds.), Silvicultural Research in a Lowland Mixed Dipterocarp Forest of East Kalimantan. Cirad, Montpellier, France, pp. 139–161.
- Sist, P., Nolan, T., Bertault, J.G., Dykstra, D.P., 1998. Harvesting intensity versus sustainability in Indonesia. For. Ecol. Manage. 108, 251–260.
- Valentine, H.T., Herman, D.A., Gove, J.H., Hollinger, D.Y., Solomon, D.S., 1998. Initializing a model stand for process-based projection. Tree Phys. 20, 393–398.
- Vanclay, J.K., 1994. Modelling Forest Growth and Yield: Applications to Mixed Tropical Forests. CAB International, Wallingford, UK.
- Vanclay, J.K., Skovsgaard, J.P., 1997. Evaluating forest growth models. Ecol. Model. 98, 1–12.
- van Gardingen, P.R., Clearwater, M.J., Nifinluri, T., Effendi, R., Rusmanto, W., Noor, M., Mason, P.A., Ingleby, K., Munro, R., 1998. Impacts of logging on the regeneration of

- lowland dipterocarp forest in Indonesia. *Comm. For. Rev.* 77, 71–82.
- van Gardingen, P.R., Foody, G.M., Curran, P.J., 1997. Scaling up: from cell to landscape. *Society for Experimental Biology, Seminar Series*, Cambridge University Press, Cambridge.
- van Gardingen, P.R., McLeish, M.J., Phillips, P.D., Dadang Fadilah, Tyrie, G., Yasman, I. 2002 Timber yield, financial and ecological analysis for the sustainable management of logged-over Dipterocarp forests in Indonesian Borneo. *For. Ecol. Manage.*, submitted for publication.