Swedish University of Agricultural Sciences
Faculty of Forestry
Uppsala, Sweden

A matrix growth model of the Swedish forest

OLA SALLNÄS
Department of Operational Efficiency

Studia Forestalia Suecica
No. 183 · 1990

ISSN 0039-3150
ISBN 91-576-4174-9

Abstract

Sallnäs, O. 1990. *A matrix growth model of the Swedish forest.* Studia Forestalia Suecica 183, 23 pp. ISSN 0039-3150, ISBN 91-576-4174-9.

An area forest matrix model was developed, intended for use as a tool for modelling forest yield in an integrated forest sector model. The model was estimated from data from the Swedish National Forest Survey. Log-linear models are used in the estimation of transition probabilities. By comparison with another growth model, the matrix model generates reasonable growth levels and growth patterns. General characteristics of the model and the matrix concept are analyzed and discussed. In general, the model is considered suitable for implementation in integrated forest sector modelling.

Key words: Forest yield, Markov model, estimation methods, log-linear models, survey data.

Ola Sallnäs, Department of Operational Efficiency, Swedish University of Agricultural Sciences, S-770 73 Garpenberg, Sweden.

Contents

MS. received 26 May 1989
MS. accepted 4 December 1989

016.93 Tofters tryckeri ab, Östervåla 1990

Introduction

In Sweden, as in other countries in which the forest sector is of prime importance to the economy, increasing interest is being shown in integrated forest sector analysis. In several countries, formal forest sector models are being developed to serve as tools for the analysis of policy. In developing these models, which attempt to encompass the forest as well as forest industry and the product market, pioneer work was done by Randers, Stenberg & Kalgraf (1978) as regards the SOS-model and by Adams & Haynes (1980) with the TAMM model. Other efforts along these lines are exemplified by Kallio, Propoi & Seppälä (1980), by Nilsson's (1980) industrially focussed model, by Lönnstedt's (1986) fairly aggregated regional model, and the trade-focussed GTM model developed at IIASA (Kallio, Dykstra & Binkley, 1987).

One essential component in a forest sector model is a module which projects the forest state, and thus serves as a basis for forest management decisions. Since the early 1970s, two major models for timber assessment studies at the regional level have been developed in Sweden. The first model was used by a government commission on forest policy (SOU, 1978). The second is the "HUGIN-system" (Bengtsson, 1981). Both are based on data from the Swedish National Forest Survey (NFS), and have been used for generating options on which government forest policy can be based. These two simulation models, HUGIN in particular, were constructed to simulate the outcome of different management programmes. Their design allows management programmes to be formulated in detail. Because of their size and computational properties, large forest models such as HUGIN are not well suited to incorporation into integrated forest sector systems. Their wealth of detail, and their lengthy running times, make them rather poorly suited for this purpose. Another of their characteristics is that the growth functions used in them (in the case of HUGIN; Ekö 1985, Söderberg, 1986) often possess a mathematical structure which makes them difficult to incorporate into an optimizing environment.

On the basis of this outline of developments in forest and forest sector modelling, it is possible to identify a need for a forest projection tool that

- can be used at a regional level,
- can be quickly and easily handled in the computer,
- can be associated with a representative description of the forest region under study,
- is sufficiently differentiated to depict the dynamics of the forest in a way which makes the interaction with a forest industrial model meaningful.

The matrix model concept is one interesting modelling approach to the objectives stated above.

The aim of this study is to develop a matrix model that could be incorporated into an integrated forest sector model. An area matrix model consists basically of three parts: (1) a matrix of forest areas, expressing the state of the forest, (2) a set of transition probabilities which, under different treatments, governs the transition of areas between the elements of the state matrix, and (3) a set of activities. This report focusses on the development of the state and transition matrices, while the question of how to derive the activity pattern is not addressed. Furthermore, the emphasis is almost entirely on the modelling of established forests. Modelling of young forest is only briefly discussed. The development of the model is presented in a series of steps:

a. The formal model is defined in general terms, together with some preset restrictions.
b. The choice of variables and the classification scheme which defines the state matrix are dealt with.
c. The data set from the Swedish National Forest Survey (NFS), used for estimating the model, is presented.
d. The methods for estimating transition probabilities are presented.

The paper concludes with a general discussion of the concept, in which the model is evaluated against other growth functions estimated from data from the NFS. In addition, some basic characteristics of the matrix model are presented and the results are discussed.

Materials and methods

Matrix models

Three major groups of forest matrix models may be found in the literature. By far the largest group encompasses models built with the single tree as the basic entity and diameter as the state-defining variable. These models are often used for modelling the development of uneven-aged or selection forests. Usher (1966, 1969, 1979) was perhaps the first to pursue the concept in a forest context. Bruner & Moser (1973) addressed the question of predicting diameter distributions. Rorres (1978) used linear programming to seek the optimal harvesting policy in an "Usher-like" matrix model. Buongiorno & Michie (1980) attempted to deal with some of the problems of exponential growth inherent in the earlier models. In particular, they dealt with the problem of modelling ingrowth. Kallio et al. (1980) used a matrix model based on the single tree, in which the states were defined by age and species, to model the forest in a forest sector model. Although the basic entity in their model was area, this area was associated with the single tree, hence the model resembles diameter-based models. Houllier (1986) made a study in which the design of forest surveys was related to the problem of constructing dynamic models based on survey data. Haight & Getz (1987) developed a diameter matrix model which was compared with associated single-tree growth functions. Leps & Vacek (1986) used a matrix model to investigate the development of a tree population with respect to vitality classes, for a situation involving air pollution.

In the second group of matrix models are models concerned with forest succession in multi-species forests. Chief interest is attached to the long-term development of species, and to size distribution in the forest studied. Many of the models in this group are "diameter-type" models, and since the objective is to analyse succession, questions concerning in-growth and species change are of prime importance. The models of Horn (1975), Barden (1981) and Bellefleur (1981) represent this group.

Area matrix models, considered here as models in which the basic entity is forest area and in which the states are defined by variables related to area, make up the third group, which occur more sparsely in the literature. One early application was presented by Hool (1966), in which an area matrix model was incorporated into an optimizing overall structure by means of a dynamic programming algorithm. Hool's model was further developed by Lembersky & Johnson (1975). Vaux (1971) outlined a basic structure for a simple model, and Kouba (1977) used an area-based matrix model for discussing the concept of the normal forest.

Common criticisms directed against forest matrix models concern the assumption of stationarity of the process (Binkley, 1980; Roberts & Hruska, 1986) and the restriction of projections to periods that are integer multiples of the growth period implicit in data used for estimating the model (Harrison & Michie, 1985). Manders (1987) suggests a procedure for testing the assumption of stationarity. The problems of assessing the effects of errors in input and parameters and of measuring uncertainties in short-term projections, were analysed by Peden, Williams & Frayer (1973). Williams (1978) later suggested possible improvements to the model. Vandermeer (1978) and Manders (1987) discussed the determination of the category size in matrix models, with particular reference to errors to be expected when estimating transition probabilities. General features of matrix models are discussed by Enright & Ogden (1979), Rottier (1984), Houllier (1986) and Manders (1987).

The model

In what follows, the general formulation of the area matrix model, its basic structure and the descriptive variables for the forest are discussed. An early prototype of the model was presented in Sallnäs, Hägglund & Eriksson (1985). Throughout this paper, a superscript index is used for denoting time, while subscript indices denote cell, state or activity.

Given a set of states S, a set of activities A and a set of transition probabilities P, the area matrix model is formulated

$$x_j^{t+1} = \sum_{i,j} x_i^t \cdot p_{ij}(k^t, k^{t-1}) \cdot a_{ik}^t \quad \text{all } i,j \in S,\ k \in A \quad (1)$$

$$\sum_k a_{ik}^t = 1 \quad \text{all } i \in S,\ k \in \bar{A} \quad (2)$$

where x_i^t is the area residing in state $i \in S$ at time t, a_{ik}^t is the fraction of the area in state i that is subject to activity $k \in A$ at time t, and $p_{ij}(k^t, k^{t-1}) \in P$ is the probability for an area residing in state i at time t to be found in state j at time $t+1$ if subjected to activity k^t at time t and to activity k^{t-1} at time $t-1$. The time step is set to five years. Three activities are allowed for in the model—thinning, final felling and no treatment.

The states in S constituting the basic forest description are defined by the variables:

- geographical region
- owner category
- site quality
- species composition
- age
- volume.

The first three of these state variables refer to the site, while the others refer to the growing stock on the site. Most yield models recently developed in Sweden use, among others, the variables region, site, age and species composition as explanatory variables (see, e.g. Eriksson, 1976; Ekö, 1985; Agestam, 1985 and Tham, 1988). In these models basal area is, in addition, used as an independent variable. In the present study, volume was, however, chosen. The variable ownership category has been introduced, since growth differences which reflect different management history may be embodied in this variable (see, e.g. Attebring, 1985; Kempe, 1980 and SOU, 1981).

The model outlined is a second-order model, in the sense that the transitions depend, not only on the activities in the present period, but also on those in the previous period. In the case of thinnings, this feature is of interest, since any thinning effect may be expected to last for more than one period. To preserve this property, while simplifying the model, it was converted to one which, from the activity point of view, is a first order model, by the introduction of a new variable, viz. thinning status. Status I indicates forests not thinned during the previous period, while status II indicates for forests thinned during the previous period. When the thinning status variable is included in the set of variables that span the state set S, the probabilities in equation (2.1) may be written $p_{ij}(k)$, or for ease of notation, p_{ijk}, where k denotes the activity applied in time t.

The three defined treatments, and the definition of the state matrix, span the theoretical set of possible transition paths. However, to simplify the model, it was decided to restrict the possible transitions. Transitions corresponding to volume growth are restricted to the augmentation of zero, one or two volume classes during one time step. The thinning activity is expressed by a reduction of the volume class by one, which takes place before growth. This implies that in the case of thinning, the possible compound transitions are that the volume class is reduced by one, remains unaltered or is increased by one. Furthermore, it was decided that the thinning treatment is not permissible in thinning status II, i.e. in forests thinned during the preceding period. Age transitions are governed by the assumption of even distribution of areas within each age class. The permitted transition paths for established forests are summarised in Table 1.

Young forests, in this study defined as bare land or forests with an average height of less than six metres, are described only by the variables region, ownership, site and age. Not until the young forest areas enter the set of established forest are they associated with a volume class and a species composition.

Classification

With a defined classification scheme and a data set of forest entities (plots, stands, etc.) describing the state for every entity at two subsequent times, a first set of state and transition matrices can be established. The state matrix contains the number of entities residing in each defined state, while the transition matrix gives the number of entities which, during the implicit five-year period, progress from one state to another. Here the classification is discussed, and in particular, the intervals chosen for the different variables. The crucial point of determining the intervals for the volume variable is dealt with in some detail.

Since the model is based on a discrete set of states, a sequence of intervals over which the data can be classified must be defined for every variable. The number of intervals affects both the estimation procedure and the computational properties of the model. A large number of intervals gives rise to large matrices and consequently, to a large number of para-

Table 1. *Possible transition paths in the model*

State at time t			Thinn. status[1]	Activity	State at time $t+1$			Thinn. status
Age	Spec.	Vol.			Age	Spec.	Vol.	
i	j	k	I	none	$i,i+1$	j	$k,k+1,k+2$	I
			II	none	$i,i+1$	j	$k,k+1,k+2$	I
			I	thinn.	$i,i+1$	j	$k-1,k,k+1$	II
			I	final	– young forest –			

[1] Thinning status II denotes forests that have been thinned in the previous period, and status I denotes forests ot thinned in the previous period.

meters to be estimated. Another consequence of large matrices is long running times. However, it is essential to depict the forest in such a way that a sufficient differentiation of growth patterns, as well as of management programmes, is possible.

The top-level classification to be made is the separation of young forests from established forests. The limit was set at an average height of six metres, a choice that may be compared with the eight-metre limit used in the HUGIN-system (Hägglund, 1981*a*). In the remainder of this section, and the entire section concerned with estimation (p. 9), established forests alone will be dealt with. Young forests are separately discussed (p. 13).

Intervals for the volume variable

The volume variable has a somewhat different status as compared with the other variables. All other variables describe state, but where dynamic behaviour is concerned, their main function is to separate different volume growth patterns. Volume growth is the core of the model, and can be recorded only as a difference in the area distribution by volume classes over a time interval. Therefore, it is essential to define the volume classes in such a way that growth is correctly depicted even in a single-period perspective.

Given the seven-dimensional classification matrix, denote by i the cell index with respect to volume class, and by j the compound cell index associated with all other variables. If the number of volume classes is assumed to be N, the growth g^*_{ij} expected in the model for a unit area in cell ij can be expressed as

$$g^*_{ij} = \sum_{m=1}^{N} p_{ijm}(k) \cdot d_{im} \qquad (3)$$

where $p_{ijm}(k)$ denotes the probability for the area in class ij to be found in volume class m one time period later under treatment k, and d_{im} the difference in volume between volume class i and m. The unit area has a "true" growth, g_{ij} and a minimum demand on a model of this kind is that the relation

$$\sum_{ij} g^*_{ij} = \sum_{ij} g_{ij} \qquad (4)$$

obtains. Howeve a more strict requirement would be that the relation should obtain at class level, i.e.

$$g^*_{ij} = g_{ij} \qquad \text{all } i,j \qquad (5)$$

If the volume growth g of a unit in class ij with standing volume v is estimated by a function $g(v) = \exp(f(v))$, and the residuals to the function $f(.)$ are assumed to be normally distributed, the growth for a unit can be expressed as

$$ln\ g(v) = f(v) + s \cdot e \qquad (6)$$

where s is the standard deviation of the residuals and $e \in N(0,1)$. If the deviation about the function $f(v)$ is regarded as a variation in growth for a unit with volume v, and we set

$$b_i(v) = 1/s\ (ln(v_i - v) - f(v)) \qquad (7)$$

with v_i as the upper limit of volume class i, the probability for the unit with initial volume v to grow out of the volume class may be expressed as

$$P(v + \exp(f(v) + s \cdot e)) > v_i) = 1 - \Phi(b_i(v)) \qquad (8)$$

where Φ is the cumulative normal density function. Consequently, considered over the entire class, which is assumed to contain a uniformly distributed forest area, the probability for an arbitrary unit in volume class i to grow out of the class is

$$P^*_i(v) = 1/(v_i - v_{i-1}) \cdot \int_{vi-1}^{vi} (1 - \Phi(b_i(v))\ dv \qquad (9)$$

In the case of no treatment, the $p_{ijm}(k)$'s, in expression (1), can be expected to be small for $m < i$ and $m > i+1$ and consequently g^*_{ij} may be approximated by $p_{ijn} \cdot d_{in}$ where $n = i+1$. Relation (5) then becomes

$$p_{ijn} \cdot d_{in} = g_{ij} \qquad (10)$$

Setting

$$g_{ij} = 1/(v_i - v_{i-1}) \cdot \int_{vi-1}^{vi} \exp(f(v))\ dv \qquad (11)$$

substituting P^*_i for P_{ijn} and recognising that $d_{in} = (v_{i+1} - v_{i-1})/2$, the difference between the means of volume classes i and $i+1$, makes it possible to generate a sequence of class limits once v_0 and v_1 are fixed. By means of this procedure, with functions $f(.)$ (see Appendix 1) estimated for different forest types, it was possible to create volume class sequences which, in a broad sense, accord with the average growth of the plots.

Intervals for other variables

Geographical region: Four regions were distinguished (Fig. 1). These coincide with regions used by the

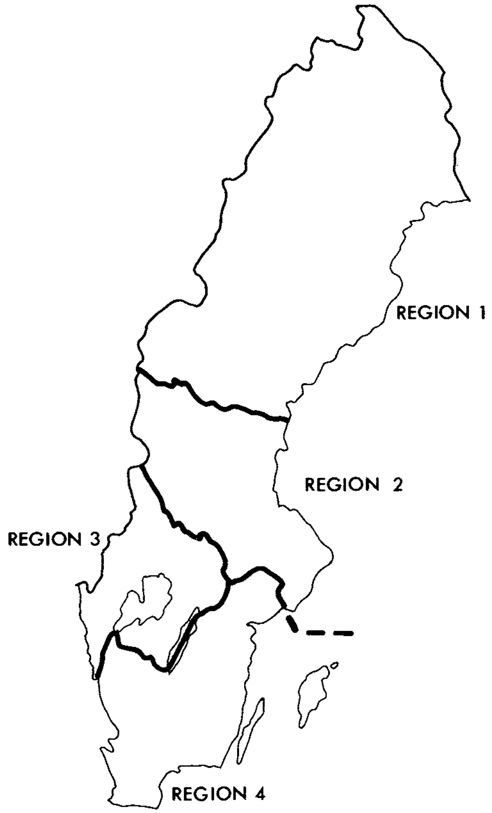

Fig. 1. The regions used in the matrix model.

government commission of enquiry into future forest policy (SOU 1981). *Ownership category:* Two ownership categories were distinguished in the model, viz. non-industrial private forest owners and others. *Site quality:* Site quality, expressed as potential mean annual yield, was used to distinguish different growth conditions. Four classes were used, the definition of which differed between regions. In regions 1 and 2, site class 1 was assigned to high-altitude forests. *Species composition:* Three classes were distinguished, viz. coniferous forests dominated by pine and by spruce, respectively, and deciduous forests. *Age:* Six age classes were used, each comprising 20 to 40 years. The intervals differed between site classes. *Thinning status:* Two different values for the thinning status variable were used, the one denoting forest not thinned during the latest five-year period, the other denoting forest thinned during the latest period.

The exact intervals used for all variables are given in Appendix 2.

Data

The model was estimated from data collected by the Swedish National Forest Survey (NFS), which annually samples the Swedish land area using systematic stratified cluster sampling, comprising about 1,000 clusters and 12,000 sampling plots. Until 1982, plots were temporary and circular, with a radius of ten metres (Bengtsson, 1978; SLU, 1974–1982). Most of the plots are "volume-plots", on which all trees are calipered. In this study, all such plots situated on forest land in the surveys from the years 1974 to 1982 were used. In total, there were about 100,000 plots. Some variables were assessed from the "20-metre plot", i.e. an imaginary plot with a radius of 20 metres, which has the same centre as the ordinary ten-metre plot.

NFS-variables and simulations

For each plot in the NFS, numerous data are collected, of which only a small number have been used in this study. Some of these data deserve further explanation. Area characteristics, such as geographical situation, altitude and ownership category, are recorded for every plot. Site index, expressed as estimated dominant height at 100 years of age, is assessed and later converted to potential mean annual yield. For every plot an age-class, relating to the 20-metre plot, is assessed. If the age is less than 40 years, an exact age is recorded. On the plots, the height of a number of sample trees is measured and they are bored for estimation of increment and volume. By a standard procedure (Holm, Hägglund & Mårtensson, 1979) each calipered tree is associated with a sample tree, thus assigning volume and growth to all trees, and consequently to all volume-plots. The data are differentiated by species. In this study, these calculated growth figures were used without adjustment for climatic variation. Natural mortality is assessed for every species, enabling the volume of losses by mortality to be simulated on the plots in connection with the other simulations. Recent treatments, together with the estimated time period in which they took place, are recorded. However, whether or not the plot has been fertilised cannot be determined (Bengtsson & Sandewall, 1978). It should, however, be recognised that the data set relates to forests in which the area fertilised annually has averaged some 150 000 hectares.

Preprocessing of the material

Before the estimation of the model, the basic data were preprocessed. The state, five years before measurement, as expressed by the variables age, volume and species composition, was assessed for every plot. The volume of species s on plot i five years before measurement, v_{is}^{-5} was calculated according to

$$v_{is}^{-5} = v_{is}^{0} - g_{is} + m_{is}$$

where v_{is}^{0}, g_{is} and m_{is} are respectively, the recorded volume and the recorded growth and mortality during the period. In this connexion, mortality was assessed directly from recorded mortality on the individual plot. It should be noted that if the plot had been thinned during the five-year-period preceding the survey, growth includes only the growth of the remaining trees. In its turn, this implies that if thinning had been carried out fewer than five years before the survey, growth covers unthinned as well as thinned conditions. Previous height was calculated, to make it possible to decide whether or not the plot should be regarded as belonging to the set "young forests" (average height less than six metres) five years before measurement. The calculations were performed using a simple relation which states that the quotient between subsequent heights equals the third root of the corresponding volume quotient.

Growth level

Due to weather conditions, the overall growth of the forest varies substantially from one year to another. These variations were analysed, to provide a picture of the relative growth level implicit in the data set used in the present study.

Growth was simulated by means of the recorded growth (obtained from increment cores) of sample trees. Sample trees are recorded by the NFS every year, but not all trees were in fact used for growth and volume simulations. Thus in some cases, growth for a specific year was simulated using increment cores from sample trees of another year.

Table 2 shows the year of recording for sample trees used for the growth estimates of different survey years. The recorded growth for sample trees from an individual year is deduced from the five outer annual rings. If the occurrences of the different annual rings in the data set are summed, the sums may be regarded as weights expressing the relative importance, to the total growth level, of each annual ring. From annual indices for different species and regions (Bengtsson & Wulff, 1987) weighted averages were computed (Table 3). They may be regarded as rough estimates of the growth levels inherent in the data set.

Table 2. *The relation between year of survey and year of recording of sample trees used for growth estimates*

Survey year	1974	75	76	77	78	79	80	81	82
Sample trees from year(s)	74/75	75	77	77	79	79	81	81	81

Table 3. *Annual indices for the different annual rings in the data set*

		Annual indices			
Annual ring	Weight	Spruce North	South	Pine North	South
69	.5	104	82	71	72
70	2	98	72	74	76
71	2	91	71	75	80
72	4	101	101	88	102
73	4	104	108	100	102
74	5.5	92	103	105	114
75	4	80	98	125	114
76	7	101	66	124	87
77	5	96	77	117	80
78	5	113	101	103	100
79	3	104	120	108	105
80	3	118	136	122	109
Weighted mean		100.5	94.3	106.7	98.4

Results

Estimation of transition probabilities

With the classification scheme defined above (p. 5) and the preprocessed data set giving the state of every plot both at the time of survey and five years before survey, a first transition matrix was established. The aim of this section is to outline methods for estimating the transition probabilities for established forests from this basic matrix (transitions from young to established forest are dealt with below). First, the non-volume transitions are briefly dealt with.

However, the primary question at issue here is the estimation of volume transitions, the main problem being to discover a method that yields estimates even for those parts of the matrix in which the number of observations is low or nil. Log-linear models, which form the basis of the estimation procedure employed, are outlined. They are used for testing which variables to use in the estimation, which is carried out as a stepwise procedure supported by a sequence of increasingly aggregated log-linear models.

The problem of estimation has the following background: If the number of states is denoted by m, the abovementioned initial matrix is of size $m.m$ and can be written $N = (n_{ij})$, where n_{ij} is the weighted number of plots belonging to state i five years before survey and to state j at the time of survey. The weights used are compounded of the size of the (part of the) plot and the sampling probability. Now the matrix of transition probabilities $P = (p_{ij})$ could be estimated by the straightforward Maximum-Likelihood estimate

$$p_{ij} = n_{ij}/\sum_j n_{ij}$$

However, in the m-dimensional state matrix, a large number of cells have none or very few observations (see Table 4). In these cases there would be no estimates or very poor ones. This situation was approached in two ways. Only volume transitions were in fact estimated (see below), thus limiting the number of parameters to be estimated, and log-linear models were used to test for the aggregation level in the estimates.

Non-volume transitions

In the matrix N, the transitions between different species composition groups are rather few in number. It should, however, be noted that in the data set it was not possible to judge whether or not a plot had crossed a species boundary when thinned, since the removed volume is not recorded. To limit the number of parameters to be estimated, it was decided to restrict the transition to take place inside the original species group. Age-transitions are depicted by transition rates equalling the quotient between the projection period of five years and the age class width. Transitions between thinning status classes are guided solely by the activity undertaken; thus if an area is thinned, it progresses to thinning status II and for the next period it returns to status I. This way of specifying the model means that all non-volume transitions inside the set of established forests are of an *a priori* nature (Table 5).

Table 5. *The number of observations in the species groups at time of recording (t) and five years earlier (t-5), in region 1*

		Time *t* Species group		
		1	2	3
Time *t*-5	Species group			
	1	9 857	129	36
	2	83	10 843	40
	3	65	149	1 905

Log-linear models

To test which variables to use for differentiating the growth patterns in different parts of the transition matrix, P, it was decided to fit log-linear models to the matrix of counts, N. Log-linear models are discussed in depth in Bishop, Fienberg & Holland (1975). A briefer presentation is given in Everitt (1977), which serves as a base for the following outline of loglinear models.

Table 4. *The distribution of the states in region 3 by number of observations*

Number of observations.	0	1	2	3	4	5	6	7	8	9	9+
Fraction of total number of states (%)	54	14	7	5	3	2	2	2	2	1	8

Starting with a two-dimensional matrix of observations, $A = (a_{ij})$ with the total number of observations denoted by T, the probability that an observation will fall in element a_{ij} may be expressed by

$$p_{ij} = a_{ij}/T$$

If there is no compound effect between the two variables associated with the indices i and j respectively, the probability can be expressed as

$$p_{ij} = p_{i\cdot} \cdot p_{\cdot j}$$

where the convention $p_{i\cdot} = \Sigma_j p_{ij}$ is used. Taking natural logarithms and converting from probabilities to expected counts, e_{ij}, will yield

$$ln\ e_{ij} = ln\ e_{i\cdot} + ln\ e_{\cdot j} - ln\ T$$

Following the notational convention of Bishop et al. (1975) this model can be rewritten as

$$ln\ e_{ij} = u + u_{1(i)} + u_{2(j)}$$

where

$$\begin{aligned} u &= 1/4 \cdot (ln\ e_{11} + ln\ e_{12} + ln\ e_{21} + ln\ e_{22}) & \\ u + u_{1(i)} &= 1/2 \cdot (ln\ e_{i1} + ln\ e_{i2}) & i=1,2 \\ u + u_{2(j)} &= 1/2 \cdot (ln\ e_{1j} + ln\ e_{2j}) & j=1,2 \end{aligned}$$

That is, we have a model, linear in the logarithms, consisting of three terms: one grand mean, one term associated with the first variable and one term associated with the second variable, i.e. a log-linear model based on the assumption of no interaction between the two variables. Generalizing to more than two dimensions is straightforward. In the three-variable case the analogue unsaturated model would be

$$ln\ e_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{12(ij)} + u_{13(ik)} + u_{23(jk)}$$

implying interactions between all pairs of variables, but assuming no threevariable interaction. It should be noted that "corresponding to particular hypotheses, particular sets of expected value marginal totals are constrained to be equal to the corresponding marginal totals of observed values" (Everitt, 1977). That is, to every log-linear model, based on a particular hypothesis about existing or non-existing interaction effects, there is a corresponding set of fixed marginal totals. For example to the model

$$ln\ e_{ijk} = u + u_{1(i)} + u_{2(j)} + u_{3(k)} + u_{23(jk)}$$

corresponds the set of fixed marginal totals $\{e_{i..}, e_{.j.}, e_{..k}, e_{.jk}\}$. Thus, given the matrix and the variables, a log-linear model can be defined by the set of fixed marginal totals. Denoting the three variables by a, b and c respectively, the above model could be described by the set of marginal totals $\{a00,\ 0b0,\ 00c,\ 0bc\}$, where a letter stands for a variable not summed over, while "0" indicates a summed-over variable. However, in this study we are dealing solely with hierarchial models, i.e. if a specific effect is included in the model, all lower order effects embedded in the original one are presumed to be included as well. Thus in this case the model could be described by the set $\{a00,0bc\}$. Finally, if we neglect the unnecessary zeros, the model is denoted $\{a,bc\}$. This notational convention is used throughout the remainder of this paper.

In some cases, log-linear models can be fitted via direct estimates, but it is often necessary to use an iterative algorithm. An algorithm, originally proposed by Deming and Stephan, and described in depth in Bishop et al. (1975), is used here.

Testing the outcome of the model, the expected counts, against the observed counts constitutes a way of testing the assumption about independence between the two variables. There are several options for measuring this goodness of fit. The usual χ^2 measure

$$\chi^2 = \sum_i (x_i - m_i)^2 / m_i$$

with x_i as the observed value and m_i as the fitted value, is one such option. However, the likelihood ratio

$$L^2 = 2 \cdot \sum_i x_i \cdot ln\ (x_i / m_i)$$

has some computational advantages that make it preferable. Both measures are asymptotically χ^2-distributed, and share the property that if a model A yields a measure of, say, L_A with D_A degrees of freedom and another model B, including one more interaction factor, yields L_B and D_B respectively, then the difference $L_A - L_B$ is χ^2-distributed with $D_A - D_B$ degrees of freedom (Kendall, 1975). This property makes it possible to test for the improvement in fit when additional factors are included in the model.

When computing the degrees of freedom, the number of independent parameters estimated should be subtracted from the number of cells estimated. However, the resulting figure should be adjusted by subtracting the number of elementary cells with zero estimates and adding the number of parameters that cannot be estimated due to corresponding zeros in the model defining marginal totals.

Fitting of log-linear models

The matrix of observed counts is in seven dimensions, where the variables are region, ownership, site, age, species group, volume and thinning status. It has already been stated that the transitions are restricted to take place in the subspace spanned by age, volume and thinning status. Furthermore, the transitions in the volume dimension are restricted to three classes. Thus, the N matrix can be collapsed to an $n \times 3$ matrix (n_{ij}) where i is, as before, the state five years before measurement, and $j \in (0,1,2)$ is the number of volume classes the plot gains during the transition period. Volume class definitions do vary over regions and site classes, which implies that the 16 submatrices defined by these two variables should be treated separately. All other variables, i.e. ownership, age, species group, volume and thinning status, can be used to differentiate the transition patterns, but it is also possible, when estimating the probabilities, to merge the data over one or several of these variables. What we can call the explanatory power of the variables can be tested by fitting different log-linear models to the matrices of observed counts.

Because of limitations in available computer software, the analysis could not be carried out in more than 5 dimensions. In turn this makes it necessary to deal with the testing in two subsequent steps. Before the analyses it is necessary to introduce some further notational conventions. The variables, ownership category, species, age, volume, and thinning status are abbreviated c, s, a, v, and t respectively. A new variable "outclass", defined as volume class at time of measurement minus volume class five years before measurement, is denoted o. Region and site class serve in this context as separators of submatrices, and are denoted by r and i. Three different sets of the matrices of observed counts will be used in the following:

$$A_{ri} = (o,v,a,t,s)$$
$$B_{ri} = (o,v,a,c,s)$$
$$C_{ri} = (o,v \times a,c,t,s)$$

where all variables follow the notation from above and the variable $v \times a$ is the compound variable volume class × age class. Now the log-linear models can be denoted unambiguously; for example, the model including the three-factor effect $v-a-s$ and the single factor o applied on matrix A can be written $\{o,vas\}_A$.

The first step of the analysis focusses on the set of A matrices, that is the matrices defined as (o,v,a,t,s). To test for the included variables, a number of log-linear models were fitted to these matrices. All models included the marginal sum vector, or in other words, the configuration "*vats*". This configuration fixes the marginal sums over the o, outclass, variable, thus ensuring that the original sampling scheme is preserved. The combination of the variables volume and outclass was as well kept in all models, since these two variables constitute the growth level in the different region × site sub-matrices. In Table 6 some results from fitting a number of models to the data for Region 1 are given.

Starting with the most aggregated model, including only the configuration corresponding to the two-factor effect outclass-volume class, one interaction factor in turn is added. If the inclusion of the new factor improves the fit, another factor is added. Thus the improvement in fit is evaluated by the relation between the change in L^2-measure and degrees of freedom. Normal χ^2-evaluation is used for the comparison, implying that an improvement in fit is noticeable by a decrease in the L^2-measure that clearly

Table 6. *L^2 measure/degrees of freedom for different models in different site classes; REGION 1. Matrix of observed counts defined as A=(outclass, volume, age, thinning type, species group)*

	Site class			
Model	1	2	3	4
(ov, vats)	436/319	1 126/552	840/555	997/571
(ov, oa, vats)	304/251	631/542	511/501	529/463
(ov, ot, vats)	429/238	1 125/548	811/478	987/567
(ov, os, vats)	423/315	1 097/548	837/541	985/567
(ov, oa, ot, vats)	295/236	629/538	489/438	526/459
(ov, oa, os, vats)	291/247	618/538	500/497	508/459
(ov, oa, ot, os, vats)	282/232	617/534	477/434	505/455
(ova, os, vats)	214/158	460/335	401/361	426/370
(ovs, oa, vats)	258/184	577/495	457/455	459/408
(oas, ov, vats)	279/208	599/464	480/417	490/428

exceeds the loss of degrees of freedom. It can be concluded that, besides the variable volume already given, age was the variable with the greatest explanatory power. Introducing it on the two-variable level significantly improved the fit, which was also the case with the species variable. However, no significant improvement was associated with the incorporation of the thinning type variable, nor with the introduction of three-variable effects. A best-fitting model was reached when three two-factor effects, outclass—volume,outclass—age class and outclass—species were combined.

However, the variable ownership category remains to be analysed. Therefore, in step two of the analysis the test-matrices are the *B* matrices. Table 7 shows results from the fitting of models including the ownership variable, to the data set of Region 1.

It may be noted that in site classes 1 and 2, the ownership variable had a significant impact on the two-factor level. From the analysis above it may be concluded that the best-fitting model for Region 1 is a model in which the two variable effects outclass—volume, outclass—age class and outclass—species are incorporated. In site classes 1 and 2, the two-factor effect outclass—ownership category is also included. A similar testing procedure was carried out for the other regions of Sweden. In Region 3 the results were practically identical to those of Region 1. Regions 2 and 4, however, differed from the pattern distinguished so far. Table 8, relating to the results for Region 4, is similar to Table 6 but for the addition of some models.

The figures imply that the best fitting model is the one in which the three-variable effect outclass—volume—age is combined with two-variable effects, outclass-species and outclass—thinning type. In other words, the data set can support the estimation of a more detailed model than was the case for Regions 1 and 3. Moreover, it is possible to carry out the second step of the analysis, the inclusion of the variable owner category, in a somewhat different way. Since the best fitting model includes the three-factor effect outclass—volume—age, the variables volume and age can be merged to one compound variable. Thus in the second step, the *C* matrices were used for Region 4. Now, since we are working with a compound "second" variable—volume × age, the inclusion of the configuration "*ov×a*" in a model implies a three-factor effect outclass—volume—age. In this region it is noticeable that in site classes 1 and 4 there were significant improvements in fit when the ownership category variable was incorporated (Table

Table 7. L^2 *measure/degrees of freedom for different models in different site classes, REGION 1. Matrix of observed counts defined as B=(outclass, volume, age, ownership category, species group)*

	Site class			
Model	1	2	3	4
(*ov,oa,os,vacs*)	379/335	601/502	475/419	428/373
(*ov,oa,oc,vacs*)	360/333	590/500	474/417	426/371
(*ovc,oa,os,vacs*)	345/318	571/474	448/392	397/340
(*oac,ov,os,vacs*)	363/309	583/449	465/338	420/363

Table 8. L^2 *measure/degrees of freedom for different models in different site classes; REGION 4. Matrix of observed counts defined as A=(outclass, volume, age, thinning type, species group)*

	Site class			
Model	1	2	3	4
(*ov,vats*)	2 042/597	1 580/547	1 424/464	1 635/452
(*ov,oa,vats*)	974/492	1 134/444	1 140/377	1 203/351
(*ov,ot,vats*)	2 009/593	1 520/543	1 386/460	1 592/448
(*ov,os,vats*)	2 030/593	1 529/543	1 248/460	1 145/448
(*ov,oa,ot,vats*)	946/488	1 081/440	1 118/373	1 176/347
(*ov,oa,os,vats*)	959/488	1 068/440	948/373	834/347
(*ov,oa,ot,os,vats*)	930/484	1 011/436	925/369	828/343
(*ova,os,vats*)	811/387	925/337	778/294	674/289
(*ova,ot,os,vats*)	778/383	861/333	755/290	667/285
(*ovs,oa,vats*)	885/440	989/377	852/320	749/293
(*oas,ov,vats*)	920/436	1 035/395	908/332	774/309

Table 9. *L^2 measure/degrees of freedom for different models in different site classes; REGION 4. Matrix of observed counts defined as C=(outclass, volume · age, ownership category, thinning type, species group)*

	Site class			
Model	1	2	3	4
$(ov \times a, ot, os, vats)$	985/467	1 131/379	918/259	813/270
$(ov \times a, oc, ot, os, vats)$	967/465	1 130/377	916/257	806/268

9). However, because there were rather few counts in these site classes, it was decided not to include the ownership category variable.

In the case of Region 2, the situation is more difficult. The first stage of the testing procedure showed the best-fitting model to be one with two-variable effects outclass—volume, outclass—age, outclass—thinning type, as well as outclass-species, included. Since software limitations restricted the analysis to five dimensions and the absence of three-factor effects precluded a merging of variables, it is not possible to test the interaction of thinning types and ownership category. However, the ownership variable showed a very slight effect on the fit when tested without disaggregation into different thinning types. Hence it was decided not to include the ownership variable in the "best-fitting model".

The conclusions of the testing procedure described are summarised in Table 10.

Estimation of volume transitions

In the previous section, best-fitting log-linear models were established. This testing was carried out with models including the configuration *"vats"*, which conserves the original sampling pattern. If it is recalled that the configuration *"vats"* stands for the marginal sums over the outclass variable, it is clear that models including this configuration will not yield estimates in cells which correspond to zero entries in these marginal sums. In order to give estimates in these cells as well, models excluding these configurations were fitted to the observations. However, the "best-fitting" models include configurations that are not all non-zero, and consequently they will not, even in the unrestricted form, give estimates in all cells. Furthermore, models do give estimates in cells where the estimate can be expected to be poor owing to a very small number of observations in the particular entry in an associated marginal sum. Hence, a limit of five deduced observations in the individual cell was used to determine which model to use. This can be dealt with by employing a sequence of increasingly aggregated models. The most aggregated model that can be used is the model defined by the sole configuration "*o*", i.e. the estimates are taken as the average in a certain outclass over all volume classes, species, ages, etc. When applying this method, the sequence of models was determined from the results of the model testing in the previous section. In Table 11 the sequence of models used is given, as also the number of states that in subsequent steps could not be given estimates. It is clear that, in certain regions and site classes, a considerable number of estimates must be taken from aggregated models.

Table 10. *The best-fitting model for site-classes in regions. Variables are abbreviated according to outclass = o, volume class = v, age class = a, owner category = c, thinning type = t, species composition = s. Matrices are defined A = (o,v,a,t,s) and B = (o,v,a,c,s)*

Region	Site	Model
1	1 and 2	$(ov,oa,oc,os,vacs)_B$
1	3 and 4	$(ov,oa,os,vacs)_B$
2	1–4	$(ov,oa,ot,os,vats)_A$
3	1 and 2	$(ov,oa,oc,os,vacs)_B$
3	3 and 4	$(ov,oa,os,vacs)_B$
4	1–4	$(ova,ot,os,vats)_A$

Young forests

Young forests are classified according to four variables only, viz. region, ownership, site quality and age. The intervals for the first three of these coincide, of course, with those chosen for established forests. The age classification is, however, unique for the young forests. A five-year interval was chosen to correspond to the calculation time step in the model. Eleven five-year age classes were combined with a class defined as bare forest land, to form in total 384 young forest classes over the four site quality classes and two ownership categories in four regions. Probabilities relating to transitions from young forest to the set of established forests were estimated from the data set by a straightforward Maximum-Likelihood estimate. Let y_i be the number of plots that resided in young forest state i five years before measurement, and that remain in the set of young forest at time of

Table 11. *The models used for estimation and the number of states which could not be given estimates, for different regions and site classes. (Total number of states in every submatrix is 720.*

	Site class			
Model	1	2	3	4
Region 1				
(ov,oa,oc,os)B	224	104	–	–
(ov,oa,os)B	186	80	98	100
(ov,oa)B	101	54	56	82
(ov)B	0	35	46	24
(o)B	0	0	0	
Region 2				
(ov,oa,ot,os)A	462	365	328	307
(ov,oa,os)A	398	262	227	223
(ov,oa)A	371	120	115	187
(ov)A	349	60	92	152
(o)A	0	0	0	0
Region 3				
(ov,oa,oc,os)B	238	211	–	–
(ov,oa,os)B	233	208	202	246
(ov,oa)B	133	140	158	200
(ov)B	83	96	108	140
(o)B	0	0	0	0
Region 4				
(ova,ot,os)A	250	267	306	339
(ov,oa,ot,os)A	211	219	257	296
(ov,oa,os)A	182	205	238	277
(ov,oa)A	123	181	213	221
(ov)A	35	102	147	183
(o)A	0	0	0	0

measurement, and x_{ij} the number of plots that resided in the young forest state i five years before measurement and are in state j of the established forest at the time of measurement. The probability of moving from young forest state i to state j in the established forest p_{ij} is then given by

$$p_{ij} = x_{ij}/(\textstyle\sum_j x_{ij}+y_i)$$

and the probability of moving from the young forest state i to $i+1$, $u_{i,i+1}$, by

$$u_{i,i+1} = 1-\textstyle\sum_j p_{ij} \text{ and}$$
$$u_{0,1} = c$$

where c is exogenously given. The constant c can be regarded as an expression for regeneration quality.

Some characteristics of the model

In this section, some characteristics of the model are analysed in three steps. First, the overall initial growth level is compared to the figures given by other sources. Even if the model is primarily intended to be used for analysing forest at an aggregated level, it must be evaluated at a disaggregated level to ensure that it depicts the dynamics of different forest types correctly Thus, in the second step the growth of different forest types, according to the matrix model, is compared with the outcome of a well-established Swedish growth model. Finally, the growth dynamics of the model are illustrated by examples of the development of two specific forest types over age.

Growth level

The present growth level of the Swedish forests is known through the figures published by the National Forest Survey (cf. Skogsdata, 82–87). The simulations carried out with the HUGIN-system in the latest national timber assessment study (AVB-85), could also be used in a comparison (Bengtsson, 1986). In Table 12, the estimated growth levels of these systems are compared with that of the matrix model.

The comparison should be interpreted with caution, since the various growth figures relate to different management programmes. In addition, the NFS

Table 12. *Growth level in the Swedish forests according to different sources (m^3 per hectare and year)*

	Region					
	1	1'	2	2'	3	4
NFS 75–79[1]	2.6		3.7		4.6	5.5
AVB-85[2]		2.4		3.5	4.4	5.3
Matrix model[3]	2.2	2.3	3.4	3.5	4.0	4.9

[1] Skogsdata 82. Recorded gross growth incl. growth of harvested trees, Survey data from the years 75–79.
[2] Net growth. Regions 1' and 2' refer to forest land in regions 1 and 2 but excluding high altitude forests.
[3] Net growth. Regions 1' and 2' refer to the site classes 2–4 in the regions respectively, which constitute roughly the same areas as in the AVB case.

figures relate to gross growth recorded during the years 1975–79. The AVB growth figure was deduced from the simulated fellings during the period 1980–1990 and from the difference in standing volume between 1980 and 1990, while the figure from the matrix model is deduced from a single-period simulation, the forest state being depicted by the NFS-data used in this study.

In order to study the matrix model for different forest types, it was compared with the growth model of Ekö (1985) on a site/species level. From NFS-data (1973–1977), Ekö estimated growth functions for the Swedish forests. The two models in the comparison employ different ways of formulating management programmes. Since every deduced growth level is associated with a specific management programme, it was decided to assess the maximum yield according to the two models. For the matrix model, this was done by solving the linear programming problem

$$
\begin{aligned}
\text{max.} \quad & \textstyle\sum_i \sum_k u_{ik} \cdot y_{ik} \\
\text{s.t.} \quad & \textstyle\sum_k y_{ik} = x_i \qquad (12) \\
& \textstyle\sum_{ik} y_{ik} \cdot p_{ijk} = x_j \\
& \textstyle\sum_i x_i = 1 \\
& x, y \geqslant 0
\end{aligned}
$$

where u_{ik} denotes the (volume) outcome of activity k in state i, and y_{ik} the area in state i that is treated with k; x relates to the earlier defined states, and p to the transition probabilities. The maximum yield, according to Ekö's functions, was sought under the restriction that the only thinning intensity permitted was 30 per cent of the basal area. Complete enumeration was used to solve the problem. The development of the young forest is represented here by assigning a given state to the area when entering the established forest at a fixed age. Some well known starting values, originally proposed by Hägglund (1981*b*), and later used by Ekö, were used here. Ekö's functions express basal area growth per time period, for which reason the initial forest states in these cases have been expressed in terms of age, basal area and number of stems. These values were used to establish the starting states in terms of volume, by using volume functions developed by Ekö (1985). In Table 13, the maximum yield for the two models is presented for spruce and pine forests in different combinations of region and site.

The maximum growth figures of Table 13 are each associated with a management programme. To illustrate this association, the thinning strategy of the optimal management programme, according to the Ekö model, was applied to the matrix model.

Figs. 2 and 3 show the development over age of total volume yield for the models in two different region—site—species combinations.

In Fig. 2, for a pine forest in central Sweden, the curve of the matrix model shows overall a somewhat higher growth level than the curve corresponding to the Ekö functions. This is not surprising, since the site chosen for the Ekö curve is placed rather low in the range of site classes valid for the matrix model.

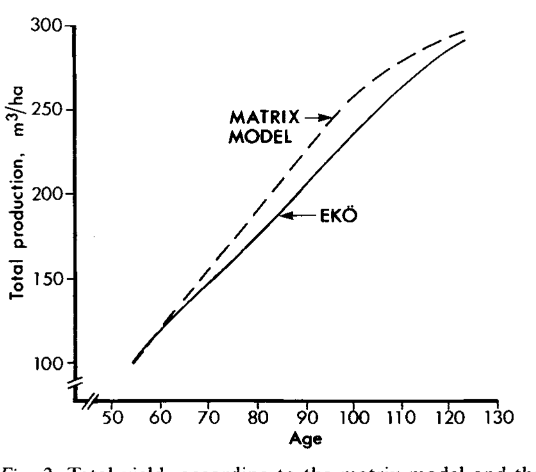

Fig. 2. Total yield, according to the matrix model and the Ekö model in a pine forest on site class 2 in region 2. The thinning programme defined maximizes volume yield in the functions of Ekö.

Table 13. *A comparison of the maximum yield according to the matrix model and the model of Ekö, for different regions, sites and species*

	Matrix model		Ekö model		Starting values			Yield	
Region species	Site class	Class-definition	Site-index	Site[1]	Age	Vol.	Basal area	Matrix	Ekö
1 pine	3	2.5 −3.5	T16	2.5	59	86	17.5	2.7	2.2
	4	>3.5	T20	3.7	41	101	19.4	4.0	3.6
2 pine	2	<4	T16	2	54	100	19.7	2.8	2.4
	3	4−6	T20	4.3	39	102	20.7	4.4	3.8
	4	>6	T28	7.7	24	108	23.1	7.1	7.6
3 pine	1	<5	T20	4.3	38	90	20.7	3.3	3.3
	2	5−7	T24	5.7	30	95	21.8	5.1	4.7
	3	7−9	T28	7.7	24	101	23.1	6.5	6.3
4 pine	1	<6	T20	4.3	38	100	23.3	4.5	3.5
	2	6−9	T28	7.7	24	100	23.3	6.1	6.1
1 spruce	3	2.5−3.5	G16	2.5	59	99	20.7	2.8	2.3
	4	>3.5	G20	3.8	47	99	20.7	3.9	3.2
2 spruce	2	<4	G16	3.1	62	101	20.7	2.8	2.9
	3	4−6	G20	4.5	47	100	20.7	4.1	4.2
	4	>6	G24	6.1	34	93	19.5	5.5	6.0
3 spruce	1	<5	G20	4.5	45	107	20.7	3.4	4.0
	2	5−7	G24	6.1	35	117	19.5	5.3	5.1
	3	7−9	G28	8.0	28	112	20.0	6.1	6.8
	4	>9	G32	9.3	23	123	21.9	8.5	9.0
4 spruce	1	<6	G20	4.3	38	122	23.3	4.5	3.5
	2	6−9	G24	6.9	35	112	19.5	5.9	6.2
	3	9−11	G32	11.3	20	120	21.9	9.7	11.1

[1] Potential annual yield, m^3 per hectare and year (PS, 1985).

However, the maximum annual yield under this management programme, according to the matrix model, is 2.6 m^3/ha, as compared with the figure of 2.8 m^3/ha in Table 13.

Fig. 3, which refers to a spruce forest on a good site, shows that the matrix model features a significantly lower growth level than the Ekö model. The maximum annual yield level of 8.5 m^3 in Table 13 is reduced to 7.5 m^3 when the thinning programme according to Ekö is applied.

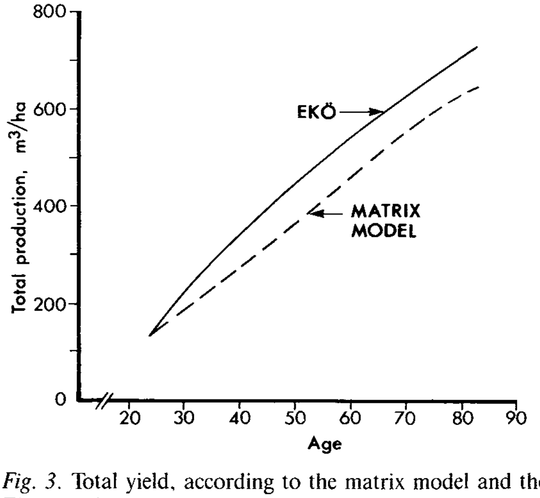

Fig. 3. Total yield, according to the matrix model and the Ekö model, in a spruce forest on site class 4 in region 3. The management programme defined maximizes volume yield in Ekö's functions.

Growth pattern

From the discussion above, it is clear that when maximum volume yield is sought, the matrix model does not favour management programmes similar to those of the Ekö model. The growth pattern of the model can be analysed through the steady-state solutions of problem (12).

The solution to one of the linear programming problems is depicted in Fig. 4 by the horizontal bars. The volume development of the forest over time, according to the optimal management programme deduced by the Ekö functions, is also shown in the figure. In this case, Ekö's functions result in a programme in which the forest is thinned three times, at the ages of 25, 35 and 45 years respectively, resulting in a low standing volume at low age, and a clear-felling at an age of about 70 years. The matrix model implies a programme in which thinning is carried out at a higher age, and final felling is postponed to an age of about 100 years.

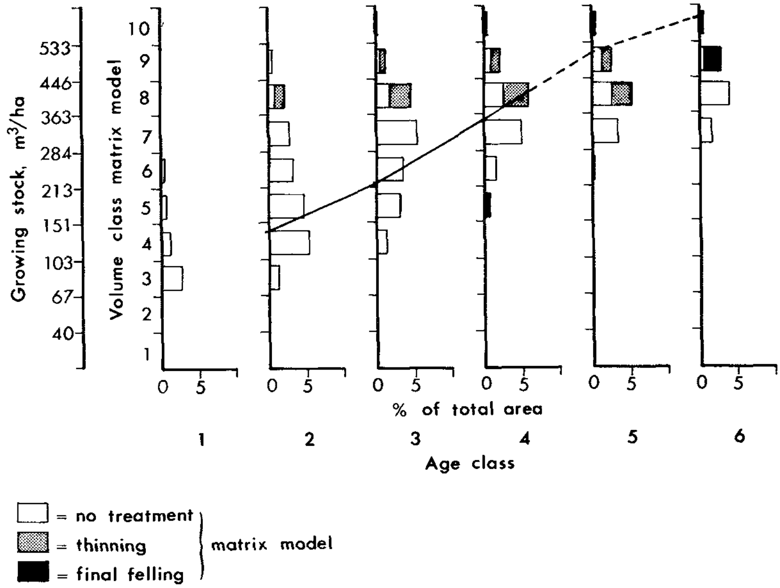

Fig. 4. The steady-state solution to (12) for a spruce forest on site class 4 in region 3 over time. Area fractions in different age and volume classes are expressed in percent. A corresponding development, according to Ekö, is indicated by the solid line (up to final felling) and by the broken line (after the imaginary final felling).

Fig. 5 is an analogue graph which refers to a pine forest on a poor site. The principal difference between the two management programmes is similar to that of the previous case. These two examples are fairly typical of the growth pattern of the matrix model. It features growth prolonged to an advanced age, and by comparison with the Ekö functions, does not take full advantage of early thinnings.

Another apparent feature of the steady-state solutions is that the forest area is distributed over several

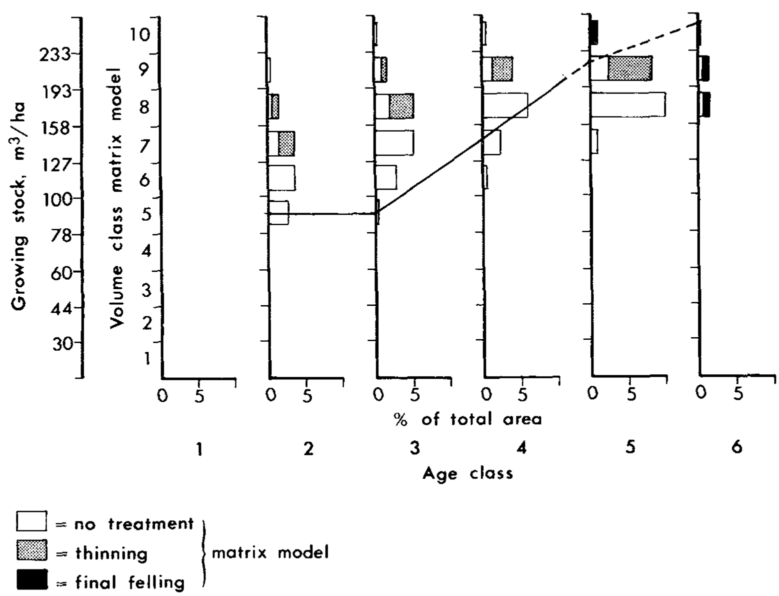

Fig. 5. The steady-state solution to (12) for a pine forest on site class 2 in region 2 over time. Area fractions in different age and volume classes are expressed in per cent. A corresponding development, according to Ekö, is indicated by the solid line (up to final felling) and by the broken line (after the imaginary final felling).

volume classes even in the lower age classes. This effect is due to the model structure, with age classes wider than the calculation period, and age transitions expressed as fractions of the area residing in a specific age class. This structure implies that some areas in an age class change not age class, but volume class, during a period, thus creating a dispersion over the volume dimension. There is also a corresponding effect over the age variable. Owing to the formulation of age transitions, some areas change age class when residing in a low volume class. The magnitude of these effects was investigated by comparing the steady-state solution in the present model, containing 6 age classes, with a corresponding solution to a model in which the forest was described by 22 age classes, each with a width of five years. The example relates to spruce forest on a good site. A management programme without thinnings, and postulating final felling at the lower limit of the sixth original age class, was applied. The expected "tilt" of the development of volume over age in the "22-class" solution compared to the "6-class" solution, is clearly evident (Table 14).

Table 14. *Average standing volumes (m^3 per hectare) at the mean age of the original age classes, according to simulations with 6 and 22 age classes respectively (Figures in brackets indicate that the volume is calculated at the lower limit of the class)*

Number of age classes used in the simulation	Age class					
	1	2	3	4	5	6
6	147	300	419	504	581	(597)
22	108	254	412	539	610	(641)

Discussion

The "growth model" presented here consists of two logical components—one yield model and one structure in which the yield model can be implemented. The yield model consists of the transition probabilities and their associated volume class scheme.

Yield model

Establishment of the yield model is concentrated to the task of classification and estimation. One important point in the classification is the definition of volume class. The choice of intervals for the volume variable is crucial to the disaggregated short-term growth level embedded in the final model. The discussion of this question in this paper is largely intuitive, and could be pursued more rigorously. Here, the importance of volume class size relates to the fact that forest areas are regarded as being represented by the mean of the class in which they reside, while the discussion of category size by Vandermeer (1978) and Manders (1987) focusses on errors expected when estimating transition probabilities. Age class intervals, which in the yield model context serve only as growth pattern separators, can be chosen *ad hoc*, but the intervals used here are not ideal. Intervals should differ between regions, not only by site class categories.

The model is based on the assumption that the variables site, species, age and volume do depict the forest in a way appropriate to the present application. Testing the influence of these variables by way of the fit of the log-linear models demonstrated that they are important. The ownership variable showed a significant effect in some areas. This may be interpreted as a result of the past application of different management regimes by different owners. In some areas a thinning effect was distinguished, noticeable in terms of a significant effect of the thinning variable in the testing of different models. Because of unreliability of the data in regions in which thinnings are few, this effect was included in the final model only in regions in which its effect was strong and logically consistent.

The choice of log-linear models as a tool in the estimation procedure should be regarded as an alternative to the straightforward Maximum Likelihood estimates commonly used (cf. Buongiorno & Michie, 1980; Mendoza & Setyarzo, 1986; Michie & McCandless, 1986; Mertens & Gennart, 1985 or Satyamurthi, 1981). The most crucial problem is how to fill in the extremes of the matrix, i.e. corners in which the number of observations is very low. By using, in the estimation procedure, the configurations identified in the testing phase, the issue can be addressed. The methods merit further investigation where log-linear models that recognise the ordered structure of some of the variables could be employed. An interesting alternative estimation procedure could be to use commonly employed stand or tree-level growth functions to generate the transition probabilities (cf. Haight & Getz, 1987, or Kaya & Buongiorno, 1987).

The NFS data are the only consistent data set covering the whole forest area of Sweden. The NFS data used in this study have, in common with most survey data based on temporary plots, two apparent drawbacks to their use in forest yield studies: no knowledge is available about fertilization of the plots, and only the mere fact that a plot has been thinned is known. The amount of wood harvested is not recorded for most plots. However, the new design of the survey, based on sampling with partial replacement, will partly solve the latter problem. Michie & Buongiorno (1984) discuss different methods that can be used for parameter estimation when remeasured sample plots are available.

In the Data section, the inherent growth level of the data set was examined. The calculations made were approximate, in that they disregarded different sampling probabilities, etc. It may nevertheless be concluded that the overall growth level seems to be rather low for spruce in southern Sweden, while it is rather high for pine in the northern part of the country. It might have been more appropriate to have used growth data adjusted for climate, instead of the unadjusted data. Furthermore, data used for estimating the development of young forest relate to forest regenerated during a period in which the quality of regeneration in Sweden was fairly low (see e.g. Kempe, 1980). The sceptical attitude to the basic assumption regarding stationarity, represented by Binkley (1980) and Roberts & Hruska (1986), is probably well founded in this case.

Implementing the yield model

The primary yield model was implemented in this study into a model possessing certain specific characteristics. In this model, the number of age classes was chosen to coincide with the number used in the estimation of the yield model. It would be possible to use a structure in which the age classes of the yield model were split, e.g. into classes with a width of five years. However, in such a case, all matrices would have been large, resulting in a large computational burden. The comparison between simulations carried out with different numbers of age classes (cf. Table 14), showed that although there were differences in volume development over age, the effects were not especially pronounced. Some of the differences between the Ekö model and the matrix model, identified in the discussion of growth pattern, probably are a consequence of the treatment of age as a series of discrete, wide classes. However, the tendency towards an improper age development in the model implies that a very skewed forest state, as in a single stand or a small forest property, would fairly soon be spread out in an inappropriate fashion.

One main structural feature of the model is its limited flexibility as regards thinning. Only one thinning activity, taking down the standing volume one class, is defined, which makes it difficult to analyse detailed management programmes. It is possible to allow additional thinning intensities, since the yield model recognises only the thinning response in the period after that in which thinning has taken place. However, the defined intensity corresponds to a harvest of 20–30 per cent of the standing volume. This level can be expected to correlate with the intensity in the thinnings carried out on the plots, hence to the intensity related to the thinning response estimated.

Fertilisation was not included in the model. This activity could, however, be incorporated fairly easily by estimating the growth response from functions. The variables used for describing the forest coincide with the most important variables in commonly used response functions (Rosvall, 1979).

The development of young forest is intentionally not dealt with in detail. The present approach means that young forest development cannot be controlled by the user. It is possible to regulate the development only by means of the coefficients controlling transitions from the bare land classes to the young forest.

In this paper, the formulation of the activity pattern of the model has not been discussed. It is stated only that the three activities allowed for in the model are thinning, final felling and no treatment, and that activities should be formulated in terms of fractions of the area in a state to be treated. The construction of these activity fractions depends, of course, on the context in which the yield model is used, but like most other matrix models, it can used in the context of both simulation and optimisation.

Concluding remarks

In the light of the evaluation of the model carried out above (p. 14), it seems reasonable to conclude that the overall growth level of the model is acceptable. Although it must be borne in mind that the growth figures relate to different management programmes, the growth of the first period compared to the NFS figures and the simulations with AVB exhibited no significant deviations, with the exception of the tendency for spruce forests in southern Sweden to deviate. This trend is further emphasised in the species/site/region comparisons with the Ekö functions. To some extent, the differences between the matrix model and the Ekö functions is explained by the circumstance that the site classes of the matrix model are somewhat wide, and that the corresponding sites chosen for the Ekö model sometimes represent extremes in the site intervals. Another explanation is given by the "poor years" inherent in the data set. When these factors are recognised,the correspondence between growth levels seems adequate. In the comparison with the Ekö functions, the matrix model was optimised. It featured management programmes that differed from those of the reference model. Thinnings were carried out comparatively late, as was the final felling. It is not surprising that models, which have different structures and are estimated by different methods, should feature such differences, even though they may be based on the same kind of data. Optimising a model, as was done in this study, may be regarded as a means of scrutinising the model for anomalies. The matrix model revealed no counter-intuitive results.

A model of the type presented in this study may be given two fundamentally different interpretations. The transitions may be regarded as fractions of areas moving from one state to others; this implies that the different transition paths result from actual differences in physical conditions, from genetic dispersion or structural properties of the forest not depicted by the descriptive variables. The other interpretation of the transitions is as probabilities, implying that the model depicts a system in which at any moment there exist several possible developments, each one associated with a probability. In this case, the development of the forest system is viewed as a stochastic process (for a discussion of the subject see, e.g., Houllier, 1986.) Furthermore, it should be noted that the model presented has the properties of a Markov-model. For this class of models, there is an extensive body of theory, available for analysing the characteristics of the model (cf. Isaacson & Johnson, 1975).

One basic question deserves attention: is the matrix model a plot level, stand level or forest level model? (For a discussion of stand vs. plot level data, see Hägglund, 1982). In the cells of the matrix reside aggregates of sample plots. The present model differs from a calculation unit model based on aggregates of plots, in that it is not based on the average values of the variables in the aggregates, but recognises that forests described in the same way do develop in different ways (or at least, that there are different probabilities associated with a number of possible developments).

As with most forest growth models, the model presented here relies on a basic assumption of stationarity. It has been stated that at least in one case, that of young forest development, this assumption is invalid. Taking into consideration the development of the genetic material used in regeneration, air pollutants, possible long-term climatic changes and other external effects, it appears plausible that the assumption of stationarity should not hold. The consequences of this will become more pronounced with time.

The properties and characteristics of the model, as discussed here, imply that it is most suitable for application to forest areas that are in some respects heterogeneous, such as large, uneven stands or forest regions, and that it should not be used for very long-term simulations.

In the Background, some desirable properties of a forest projection tool suitable for application in an integrated forest sector environment were identified. The model presented in this study can be used on a regional level and is, in consequence of the matrix structure, easily handled in a computer. It is estimated from NFS-data, which naturally can also serve as a source of continuously updated state description of the Swedish forests. Although, by comparison with most stand projection models in Sweden, the model is quite aggregated, the degree of differentiation could suffice for an integrated environment. Together with the robustness of the model, these properties imply that the model presented can serve as a part of a forest sector model.

References

Adams, D. M. & Haynes, R. W. 1980. The 1980 Softwood Timber Assessment Market Model: Structure, Projections, and Policy Simulations. Supplement to *Forest Science, Monograph 22*.

Agestam, E 1985. *En produktionsmodell för blandbestånd av tall, gran och björk i Sverige* (A growth simulator for mixed stands of pine, spruce and birch in Sweden) (Swedish University of Agricultural Sciences, Department of Forest Yield Research, Report 15, Garpenberg).

Attebring, J. 1985. *Varför ar gapet mellan tillväxt och bonitet så stort? En orienterande studie av skogarna i Småland* (Why is the gap between increment and site productivity so large? An introductory study of the forests in Småland) (Swedish University of Agricultural Sciences, Department of Forest Survey, Report 48).

Barden, L. S. 1981. Forest development in canopy gaps of a diverse hardwood forest of the southern Appalachian Mountains. *Oikos 37* 205–209.

Bellefleur, P. 1981. Markov models of forest-type secondary succession in coastal British Columbia. *Can. J. For. Res. 11*, 18–29.

Bengtsson, G. 1978. The Swedish national forest survey— *IUFRO meeting, Bucharest.*

Bengtsson, G. 1981. *Methods for Long-Term Forecasts of Timber Yields and Possible Cut Developed by the Swedish HUGIN Project.* Dept. of Forest Survey, Swedish University of Agricultural Sciences.

Bengtsson, G. 1986. *Avverkningsberäkning 1989. Sammanfattning av viktigare resultat och slutsatser.* Sveriges lantbruksuniversitet, Inst. för skogstaxering.

Bengtsson, G. & Sandewall, M. 1978. *Kartläggning av gödsling på riksskogstaxeringens provytor.* HUGIN-rapport 8. Sveriges lantbruksuniversitet, Inst. för skogstaxering.

Bengtsson, G. & Wulff, S. 1987. *Riksskogstaxeringens skogsskadeinventering år 1987 samt sammanfattning av resultaten från 1984–87.* Sveriges lantbruksuniversitet, Inst. för skogstaxering.

Binkley, C. S. 1980. Is succession in hardwood forests a stationary Markov process? *Forest Science 26* (4), 566–570.

Bishop, Y. M. M., Fienberg, S. E & Holland P. W. 1975. *Discrete Multivariate Analysis: Theory and Practice.* The Massachusetts Institute of Technology.

Bruner, H. D. & Moser, Jr, J. W. 1973. A Markov chain approach to the prediction of diameter distributions in uneven-aged forest stands. *Can. J. For. Res. 3*, 409–417.

Buongiorno, J. & Michie, B. 1980. A matrix model of uneven-aged forest management. *Forest Science 26* (4), 609–625.

Ekö, P. M. 1985. *En produktionsmodell för skog i Sverige, baserad på bestånd från riksskogstaxeringens provytor* (A growth simulator for Swedish forests, based on data from the national forest survey) (Swedish University of Agricultural Sciences, Dept. of Silviculture, Report 16).

Enright, N. & Ogden, J. 1979. Applications of transition matrix models in forest dynamics: Araucaria in Papua New Guinea and Nothofagus in New Zealand. *Australian Journal of Ecology 4*, 3–23.

Eriksson, H. 1976. *Granens produktion i Sverige* (Yield of Norway spruce in Sweden) (Royal College of Forestry, Department of Forest Yield Research, Research Notes 41). Stockholm.

Everitt, B. S. 1977. *The Analysis of Contingency Tables.* London: Chapman & Hall.

Haight, R. G. & Getz, W. M. 1987. A comparison of stage-structured and single-tree models for projecting forest stands. *Natural Resource Modeling 2* (2), 279–298.

Harrison, T. P & Michie, B. R. 1985. A generalized approach to the use of matrix growth models. *Forest Science 31* (4), 850–856.

Holm, S., Hägglund, B. & Mårtensson, A. 1979. *En metod för generalisering av Riksskogstaxeringens provträdsdata* (A method for generalization of sample tree data from the Swedish National Forest Survey) (Swedish University of Agricultural Sciences, Department of Forest Survey, Report 25).

Hool, J. N. 1966. A dynamic programmeming-Markov chain approach to forest yield control. *Forest Science, Monograph 12.*

Houllier, F. 1986. Echantillonnage et modelisation de la dynamique des peuplements forestiers. Application au cas de l'inventaire forestier national. Université Claude Bernard, Lyon.

Horn, H. S. 1975. Markovian process of forst succession. In: *Ecology and evolution of communities* (ed. M. L. Cody and J. M. Diamond), Cambridge, Mass.: Belknap Press.

Hägglund, B. 1981*a*. *Forecasting Growth and Yield in Established Forests. An outline and analysis of the outcome of a subprogramme within the HUGIN project* (Swedish University of Agricultural Sciences, Department of Forest Survey, Report 31).

Hägglund, B. 1981*b*. *Samband mellan ståndortsindex H 100 och bonitet för tall och gran i Sverige.* HUGIN-rapport 26. Sveriges lantbruksuniversitet.

Hägglund, B. 1982. Some remarks on the plot-stand problem in forest inventory. In *Statistics in Theory and Practice. Essays in Honour of Bertil Matérn ed,* B. Ranneby) (Swedish University of Agricultural Sciences, Section of Forest Biometry).

Isaacson, D. L. & Madsen, R. W. 1975. *Markov Chains—theory and application.* New York: Wiley & Sons.

Kallio, M., Propoi, A. & Seppälä, R. 1980. A Model for the Forest Sector. *WP-80-34, March 1980. IIASA, Austria.*

Kallio, M., Dykstra, D. P. & Binkley, C.S. (eds.). 1987. *The Global Forest Sector. An Analytical Perspective.* New York: John Wiley & Sons.

Kaya, I. & Buongiorno, J. 1987. Economic harvesting of uneven-aged northern hardwood stands under risk; a Markovian decision model. *Forest Science 33* (4), 889–907.

Kempe, G. 1980. *Riksskogstaxeringens återväxtinventering åren 1973–77* (Regeneration condition according to the National Forest Survey in 1973–77) (Swedish University of Agricultural Sciences, Department of Forest Survey, Report 28).

Kendall, M. 1975. *Multivariate analysis.* London: Griffin & Co.

Kouba, J. 1977. Markov chains and modelling the long-term development of the age structure and yield of forests.—proposal of a new theory of the normal forest. *Scientia Agriculturae Bohemoslovaca 9*, 179–193.

Lembersky, M. R. & Johnson, K. N. 1975. Optimal policies for managed stands: An infinite horizon Markov decision process approach. *Forest Science 21* (2), 109–122.

Leps, J. & Vacek, S. 1986. Use of transition matrices for prediction of the development of the spruce population in the Giant Mountains. In *Simulation of Systems in Biology and Medicine,* 1986, Dum techniky CSVTS, Praha.

Lönnstedt, L. 1986. A Dynamic Forest Sector Model with a Swedish Case. *Forest Science 32* (2), 377–404.

Manders, P. T. 1987. A transition matrix model of the population dynamics of the clanwilliam cedar (*Widdringtonia dedarbergensis*) in natural stands subject to fire. *Forest Ecology and Management 20*, 171–186.

Medoza, G. A. & Setyarso, A. 1986. A transition matrix forest growth model for evaluating alternative harvesting schemes in Indonesia. *Forest Ecology and Management 15*, 219–228.

Mertens, P. & Gennart, M. 1985. Gestion de la futaie jardinée. Un modèle simple de calcul des accroissements et de simulation de la yield. *Bulletin de la Société Royale Forestière de Belgique 92* (3).

Michie, B. R. & Buongiorno, J. 1984. Estimation of a matrix model of forest growth from re-measured permanent plots. *Forest Ecology and Management 8*, 127–135.

Michie, B. R. & McCandless, F. D. 1986. A matrix model of oak-hickory stand management and valuing forest land. *Forest Science 32* (3), 759–768.

Nilsson, S. 1980. *Svensk skogsnäring i omdaning. En strukturstudie med världsbankens analysmodell.* Inst för skogsteknik, Sveriges lantbruksuniversitet.

Peden, L. M., Williams, J. S. & Frayer, W. E. 1973. A markov model for stand projection. *Forest Science 19* (4), 303–314.

PS. 1985. *Praktisk Skogshandbok.* Sveriges Skogsvårdsförbund, Djursholm.

Randers, J., Stenberg, L. & Kalgraf, K. 1978. *Skognaeringen i overgangsalderen. En analyse av overgangen fra raskt voksende til tilnaermet konstant produksjonsvolum i Nordens skognaering.* J. W. Cappelens Forlag.

Roberts, M. R. & Hruska, A. J. 1986. Predicting diameter distributions: a test of the stationary Markov model. *Can. J. For. Res. 16*, 130–135.

Rorres, C. 1978. A linear programmeming approach to the optimal sustainable harvesting of a forest. *J. Environ. Mgmt 6*, 245–254.

Rosvall, O. 1979. Prognosfunktioner för beräkning av gödslingseffekter. Särtryck ur *Föreningen Skogsträdsförädling och Institutet för skogsförbättring, årsbok 1979.*

Rottier, H. 1984. A matrix model of uneven-aged forest stands. *Netherlands Journal of Agricultural Science 32*, 242–245.

Sallnäs, O. Hägglund, B. & Eriksson, L. E. 1985. *A Matrix Model of the Swedish Forest* (Swedish University of Agricultural Sciences, Department of Operational Efficiency, Working Paper 32).

Satyamurthi, K. R. 1981. Application of stochastic processes to predict forest stand dynamics. Baden-Württembergische forstliche Versuchs- und Forschungsanstalt. *Mitteilungen 91*, 127–131.

Sveriges lantbruksuniversitet, 1974–82. Inst. för Skogstaxering. Instruktion för fältarbetet vid Riksskogstaxering år 1982

SOU: 1978: 7. *Skog för framtid.* Betänkande av 1973 års skogsutredning. Bilagor. Stockholm 1978.

SOU: 1981: 81. *Skogsindustrins virkesförsörjning.* Betänkande av virkesförsörjningsutredningen. Stockholm 1981.

Skogsdata, 1982–87. 1987. Sveriges lantbruksuniversitet, Inst. för Skogstaxering. Aktuella uppgifter om de svenska skogarna från riksskogstaxeringen.

Söderberg, U. 1986. *Funktioner för skogliga produktionsprognoser – Tillväxt och formhöjd för enskilda träd av inhemska trädslag i Sverige* (Functions for forecasting to timber yields – Increment and form height for individual trees of native species in Sweden) (Swedish University of Agricultural Sciences, Section of Forest Mensuration and Management, Report 14).

Tham, Å. 1988. *Yield prediction after heavy thinning of birch in mixed stands of Norway spruce* (Picea abies *(L) Karst.*) *and birch* (Betula pendula *Roth* & Betula pubescens *Ehrh.*) (Swedish University of Agricultural Sciences, Department of Forest Yield Research, Report 23).

Usher, M. B. 1966. A matrix approach to the management of renewable resources, with special reference to selection forests. *Journal of applied ecology 3.*

Usher, M. B. 1969. A matrix model for forest management. *Biometrics 25*, 309–315.

Usher, M. B. 1979. Markovian approaches to ecological succession. *Journal of Animal Ecology 48* (2), 413–426.

Vandermeer, J. 1978. Choosing category size in a stage projection matrix. *Oecologia 32.*

Vaux, H. J. 1971. In: Folia Forestalia 101. Forecasting in forestry and timber economy. *IUFRO, Section 31, Working Group 4.*

Williams, J. S. 1978. Markov models for stand projection. In Fries, J. (ed.) *Growth Models for Long Term Forecasting of Timber Yields.* School of Forestry and Wildlife Resources, VPI, Blacksburg.

Appendix 1

Growth functions used for generating volume classes:

General model $g = a \cdot s^b \cdot \exp(c \cdot v)$

where g is growth per cent per five year period, s site expressed in potential production measured in m^3/100 hectares/year and v is growing stock (m^3/ha).

Coefficients

Region	Site class	a	b	c
1	1	2.94	0.3119	−0.0039
1	2–4	1.04	0.5790	−0.0047
2	1	6.87	0.1710	−0.0040
2	2–4	0.99	0.5709	−0.0043
3	1–4	1.03	0.5713	−0.0045
4	1–4	2.18	0.4465	−0.0042

Appendix 2

Classification scheme for the NFS-plots

VARIABLE

SPECIES COMPOSITION (% of standing volume)	1	>50% conifers, pine dominated (% of standing volume)
	2	>50% conifers, spruce dominated
	3	>50% broadleaves

REGION

		1	2	3	4
SITE CLASS m^3/ha, years potential mean increment	1	>500 m altitude	>600 m altitude	<5	<6
	2	<2,5	<4	5–7	6–9
	3	2,5–3,5	4–6	7–9	9–11
	4	>3,5	>6	>9	>11

SITE CLASS

		1	2	3	4	1	2	3	4	1	2	3	4	1	2	3	4
VOLUME CLASS (Upper limit) m^3/ha	1	25	25	30	30	30	30	35	40	30	35	40	45	35	45	50	55
	2	35	36	45	48	42	44	54	65	45	57	67	78	53	72	82	92
	3	46	49	64	70	55	60	78	99	63	85	103	123	75	107	125	142
	4	57	63	86	100	68	78	109	144	86	122	151	185	102	153	181	208
	5	70	80	113	138	84	100	147	201	114	169	213	262	136	210	249	287
	6	85	101	146	184	101	127	193	269	147	225	284	349	176	275	327	376
	7	103	125	184	237	121	158	246	345	187	289	363	442	223	348	411	469
	8	123	152	227	296	143	193	305	426	231	358	446	540	275	425	499	567
	9	145	183	273	358	168	233	367	510	280	432	533	645	332	505	591	669
	10	>145	>183	>273	>358	>168	>233	>367	>510	>280	>432	>533	>645	>332	>505	>591	>669

SITE CLASS

		1	2	3	4
AGE CLASS (Upper limit)	1	50	40	30	20
	2	70	60	50	40
	3	90	80	70	60
	4	120	100	90	80
	5	160	140	120	100
	6	>160	>140	>120	>100