STAND GROWTH MODELS: THEORY AND PRACTICE

Oscar García
Instituto Forestal
Santiago, Chile

ABSTRACT

A tutorial on growth modelling principles and methods. A system-theoretic viewpoint is adopted to establish a proper foundation for growth modelling and unify the various approaches and methods. The strengths and weaknesses of models at various levels of resolution are reviewed. There is a detailed discussion of parameter estimation. Results and experience from applying the concepts and techniques in an extensive growth modelling program in New Zealand are described as an example. Some growth patterns observed during this work are examined, including issues related to self-thinning.

Keywords: Growth models, System Theory, state-space, dynamic systems, site index, thinning, Statistics, estimation, self-thinning.

INTRODUCTION

Foresters can influence the development of a forest stand through a number of decisions and silvicultural treatments. Stand density can be controlled by choosing an initial planting density, and by thinning. Density affects total production, the incidence of competition-induced mortality, and the size and timber quality of individual trees. Timber quality can also be improved by pruning, possibly at some cost in reduced growth. Other management decisions may involve the application of fertilizers and pesticides, different planting and regeneration techniques, the development and use of genetically improved seed, and the timing of the final cut. It is clear that, because of the long production cycles and of the numerous management alternatives, the possibilities of learning from direct experimentation or through trial and error are limited. Consequently, mathematical growth models capable of predicting treatment effects are essential for rational forest management, especially in intensively managed production forests.

In general, a model is a simplified representation of some aspect of reality (not to be confused with the normative meaning of the word, something worthy of being imitated.) We continuously use models, in some form or other. There are mental models, i. e., imagined cause-effect relationships between components of some system through which we try to explain and anticipate its behavior. Models may be stated in verbal form, for example the description in words of the workings of some machine. Material models, such as scale models of buildings and airplanes, are well-known.

A mathematical model is like a verbal model, but using mathematical language. Mathematical language differs from natural language in that it is more concise and less ambiguous. This, together with the availability of pre-digested reasoning rules that can be

used mechanically, allows us to reason in more complex situations, with less effort, and with less risk of confusion. With the progress in computing, it becomes easier to handle increasingly complex models, and in fact, the use of engineering scale models has been decreasing, replaced by mathematical models that are cheaper and more flexible. Computers have become indispensable as tools in the development and use of many models. But note that *computer modelling* is analogous to *typewriter poetry*. Also, realism is not necessarily a virtue in a model, it is better to abstract just those aspects that are most relevant in each instance. In a model airplane for wind tunnel testing, its color or the name of the pilot may not be important. The manual for a video recorder does not say much about its inner workings, but it can be used to predict the effects of pressing different buttons.

We are interested here in mathematical models for forest stands. It is useful to distinguish between *models for prediction*, and *models for understanding* (Bunnell 1989). Models for understanding (e.g. physiological, process models) are mostly useful in research, to help comprehension, to collect and link previously isolated bits of knowledge, and to identify gaps where more work is needed. The benefits come from the development of the model, and not so much from its later use, if any. To be specific I shall focus on models for prediction, although the scope of many of the ideas is more general. These models are intended for management. Typical applications are in forecasting for forest planning purposes, in the comparison and evaluation of silvicultural (pruning, thinning) regimes, and in the updating of stand description databases.

The presentation is built around state-space concepts, which in a number of fields have helped to think clearly about systems that evolve in time. These are explained in the next section. Then, the foundations of growth modelling are examined from this point of view, and the various types of models are characterized. This is followed by a section on the modelling of the important height growth component, and its use for quantifying site quality. For the most part the models are presented as deterministic. It is obvious that in reality stands will deviate from the model predictions. In general, decision-makers use model results as representing a most likely course of events, and introducing "randomness" (whatever that means) has not been found very helpful in practice so far. A stochastic (random, probabilistic) component is however necessary for developing rational parameter estimation procedures. Stochastic structures and parameter estimation methods for growth modelling are discussed in some detail. Following that, the application of much of this theory to the development of a series of growth models in New Zealand is briefly described. Finally, the state-space framework and experience from the New Zealand modelling program are used in an attempt to improve our understanding of some aspects of stand development. These include, at one extreme, the growth of intensively managed, heavily thinned and pruned stands, and at the other, self-thinning in dense stands undergoing competition-induced mortality.

Although mathematical notation has been used without condescension wherever it seemed appropriate, the paper should be understandable to anyone with an elementary mathematics and forestry background.

DYNAMIC SYSTEMS

A growth model predicts future values of certain *outputs*, e. g. timber volume, given *inputs* (control variables) such as silvicultural treatments. Both inputs and outputs are functions of time. This fact, and the dependency of the outputs on the entire past history of the stand, have caused considerable difficulties and confusion in growth modelling. The way out is well-known in other fields dealing with dynamic systems, and may be called the *state-space approach* [1].

The first trick is to characterize the *state* of the system at any point in time so that given the present state the future does not depend of the past. For example, we could characterize a stand by its basal area, stocking (stems per hectare), and top height, and assume that two stands with the same values for these variables would behave in practically the same way, no matter how they happened to reach that state. "Roughly, a state of a system at any given time is the information needed to determine the behavior of the system from that time on" (Zadeh 1969). In modelling we can always describe a system in enough detail to determine its future behavior up to a desired level of accuracy; the existence of a state description is a matter of definition. From the point of view of doing away with the past, the state may be "regarded as a kind of information storage or memory or an accumulation of past causes" (Kalman, Falb and Arbib 1969), "some compact representation of the past activity of the system complete enough to allow us to predict, on the basis of inputs, exactly what the outputs will be, and also to update the state itself" (Padulo and Arbib 1974).

Let the state at a given time t be specified by a list of n numbers (*state variables*), that is, by an n-dimensional *state vector* $\mathbf{x}(t)$. The inputs and outputs are also finite-dimensional vectors [2] $\mathbf{u}(t)$ and $\mathbf{y}(t)$, respectively (more generality is possible, but usually not needed.) Then the behavior of the system is described by a *transition function*

$$\mathbf{x}(t) = \mathbf{F}[\mathbf{x}(t_0), \mathbf{U}, t - t_0] \ , \tag{1}$$

and an *output function*

$$\mathbf{y}(t) = \mathbf{g}[\mathbf{x}(t)] \ . \tag{2}$$

In words, (1) (which without vector notation would be written as a system of n equations) gives the state at any time t as a function of the state at some other time t_0, the inputs

[1] These ideas were developed most clearly in System Theory in the late 1960's and early 70's, but were already used by its predecessors, the Control Theory of the 60's and the Cybernetics of the 50's. In a specialized form they can be traced back to Mechanics, with contributions from certain areas within the Theory of Differential Equations. The most active current successor of System Theory seems to be the field of Nonlinear Dynamics.

[2] Vectors are just shorthand for a list of numbers, $\mathbf{x} = (x_1, x_2, \ldots, x_n)$. Vector functions, like (1) and (2), save even more writing: $\mathbf{y} = \mathbf{f}(\mathbf{x})$ is the same as the system of equations

$$\begin{aligned} y_1 &= f_1(x_1, x_2, \ldots, x_n) \\ y_2 &= f_2(x_1, x_2, \ldots, x_n) \\ \vdots &\quad \vdots \\ y_n &= f_n(x_1, x_2, \ldots, x_n) \end{aligned}$$

(as a function of time, denoted by $\mathbf{U}$), and the elapsed time between t_0 and t. The output function (2) gives the current outputs as a function of the current state.

A transition function must possess some obvious properties:

1. (Consistency) No change for zero elapsed time, i. e.,

$$\mathbf{F}[\mathbf{x}(t), \mathbf{U}, 0] = \mathbf{x}(t) \quad \text{for all} \quad t, \mathbf{x}(t), \mathbf{U} \; .$$

2. (Composition or semigroup property) The result of projecting the state first from t_0 to t_1, and then from t_1 to t_2, must be the same as that of the one-step projection from t_0 to t_2:

$$\mathbf{F}[\mathbf{F}[\mathbf{x}(t_0), \mathbf{U}, t_1 - t_0], \mathbf{U}, t_2 - t_1] = \mathbf{F}[\mathbf{x}(t_0), \mathbf{U}, t_2 - t_0] \; , \quad \text{for any} \quad t_0 \leq t_1 \leq t_2 \; .$$

3. (Causality) A change of state can only be influenced by inputs within the relevant time interval:

$$\mathbf{F}[\mathbf{x}(t_0), \mathbf{U}_1, t_1 - t_0] = \mathbf{F}[\mathbf{x}(t_0), \mathbf{U}_2, t_1 - t_0] \quad \text{if} \quad \mathbf{u}_1(t) = \mathbf{u}_2(t) \quad \text{for} \quad t_0 \leq t \leq t_1 \; .$$

The second trick is to exploit the fact that transition functions generated by integration of differential equations (or summation of difference equations when using discrete time) automatically satisfy these conditions. The model can then be stated as follows [3] :

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \mathbf{u}) \tag{3}$$

$$\mathbf{y} = \mathbf{g}(\mathbf{x}) \; . \tag{4}$$

In a discrete-time model $\Delta\mathbf{x}$ is substituted for $d\mathbf{x}/dt$. Integration of the *local transition function* (3) between t_0 and t gives the (*global*) transition function (1).

In the state-space approach, therefore, we avoid trying to model directly the complex relationship between inputs and outputs over time. Instead, we describe the state of the system at a point in time, and model the rate of change of state (3). The state must be such that, to a sufficient degree of approximation, future states are determined by the current state and current and future actions, (3), and any quantities of interest can be estimated from the state variables, (4).

I have taken the output function as representing the link between the state variables and other variables that are required in the applications. A somewhat different interpretation is often useful in parameter estimation problems. There, the internal state of the system may be partially or entirely hidden, and the outputs are values that can be observed. This view will be used in a later section.

[3] The derivative in the left-hand-side of (3) simply denotes the rate of change in $\mathbf{x}$. No previous knowledge of differential equations will be required and, if preferred, the reader can substitute throughout discrete time increments for the instantaneous rates.

GROWTH MODELS, RESOLUTION LEVEL

A soundly-based growth model should be capable of being stated as in (3)–(4) or its discrete-time counterpart (García 1974, 1979.) An exception are the so-called "static" models, for stands that are unthinned or subject to standardized silvicultural regimes, where the absence of input alternatives allows outputs to be modelled as fixed functions of time.

Usually it is not necessary to include control variables **u** in (3). Silvicultural treatments (thinning, pruning) normally occur at discrete points in time, causing an instantaneous change of state. We can then model the stand in between treatments as a *free* system, where the rate of change of state is a function of the current state only. Separate functions can represent the changes in state variables caused by treatments, for example the reduction in basal area resulting from thinning various numbers of trees per hectare [4].

The various kinds of growth models differ in the level of detail included in the state description. Stand-level models characterize the state of the stand by a few variables representing stand-level aggregates, such as basal area, mean diameter, volume per hectare, stems per hectare, average spacing, top height, etc. At the other extreme, distance-dependent individual-tree growth models include the location (coordinates) and diameter, and sometimes height and crown dimensions, for all the trees in a sample plot. Distance-independent individual-tree models use diameter distributions in various forms, or lists of tree sizes, but without requiring tree coordinates. Individual-tree models are likely to be needed in the more complex situations: uneven-aged and/or mixed species stands, or planting in rows or other patterns in agroforestry. For reasonably homogeneous even-aged pure stands there is a choice, and parsimony would suggest not using more state variables than necessary. If needed, size distributions can often be satisfactorily estimated from stand-level variables, that is, dealt with in the output function (García 1984).

Some potentially serious problems with individual-tree models have generally been ignored (García 1992). Stands are not simple sets of trees; tree sizes show spatial structure. Micro-site similarities make neighboring trees to be more alike than average, while competition has the opposite effect. The resulting spatial correlations cause size distributions to vary unpredictably with plot size. In particular, distributions obtained from plots will differ from those for whole stands, which are the ones required in the applications. Also, distance-dependent models have until now been based only on competition indices, ignoring micro-site effects. It seems that developing satisfactory individual-tree models might be considerably more difficult than previously thought.

Here I shall focus largely on stand-level models, for pure, even-aged stands. What would be an appropriate state vector? Consider the top height H as a one-dimensional state description. For a given site, the rate of change of H (height increment) can be modelled adequately as a function of the current H: $dH/dt = f(H), or \Delta H = f(H)$. If we were interested in predicting volume per hectare, however, this output cannot be estimated satisfactorily from H alone. We would need more state information.

[4] Formally, the transition function is split into two, one for growth transitions and another for treatment transitions. We will largely ignore the second one, and only refer to the first one as the transition function, modelling growth in-between treatments.

Let us add then the stand basal area, B, forming a two-dimensional state vector $\mathbf{x} = (H, B)$. The (local) transition function can be written as a system of two equations [5] :

$$\frac{dH}{dt} = f_1(H) \tag{5}$$

$$\frac{dB}{dt} = f_2(H, B) \ , \tag{6}$$

and the output (volume) can be given as

$$V = g(H, B) \ , \tag{7}$$

for example by a stand volume equation of the form $V = a + bBH$. This two-dimensional formulation is close to many of the classical stand-level models (Clutter et al. 1983), and it may be useful to delve in it to discuss a few issues and to illustrate basic concepts. The objective of these models is to predict volume for various thinning regimes and initial spacings.

The model equations usually include site index or some other measure of site quality. Being a constant for a particular stand, I have not indicated this variable explicitly. Equation (5) does not contain B, adopting the usual assumption that top height is not affected by stand density. For the same reason, H is not altered by treatment, and it is customary to specify H as a function of age and not through the increments.

The core of the model is (6). The basal area increment is commonly taken as a function of B and age, effectively using age as a state variable instead of H (Clutter 1963, Pienaar and Turnbull 1973). Pienaar and Turnbull (1973) make explicit the idea of the two-dimensional state space in their thinned stand hypothesis: "For a wide range of thinning regimes the growth rate in a thinned stand is identical with that of an unthinned stand of the same age and the same basal area as the thinned stand". The use of age on the right-hand-side is conceptually unsatisfactory in that, at least in the sense of elapsed time t, it does not have a physical presence (other than as a number of growth rings), and therefore should not be given a causal meaning. Actually, when foresters say *age* they often think *size*. In particular, according to the model and for a given site, top height is functionally related to age, and we can substitute H for age and obtain an equation like (6). In practice this substitution does not make any difference, but it is conceptually cleaner and facilitates thinking in more complicated situations.

As an example, let us take the basal area increment model of Clutter (1963) for site index 70, $dB/dt = B(5.55 - \ln B)/t$. There is no height model in that paper, so let us use $\ln H = 4.78 - 13.3/t$ which for site 70 approximates that in Clutter and Lenhart (1968) to within one foot over the range of their data (using their equation would be cumbersome because it cannot be solved in closed form for t.) From this, the following state-space model can be obtained:

$$\frac{dH}{dt} = 0.0752H(4.78 - \ln H)^2 \tag{8}$$

$$\frac{dB}{dt} = 0.0752B(5.55 - \ln B)(4.78 - \ln H) \ . \tag{9}$$

[5] Annual or periodic increments Δx_i may be substituted for the instantaneous growth rates dx_i/dt.

Directions and speeds of movement given by (8) and (9) are shown in Figure 1. The lengths of the line segments represent one year's growth. Technically, these local transition equations define a two-dimensional vector field over the whole state space H–B (although this does not matter for points such as those toward the upper-left, which are not reachable in practice.) A stand left for itself anywhere in the plane will move as indicated by the line segments. Take a stand 13 feet high, with a basal area of 12 square feet per acre at 6 years of age. If left unthinned, it would follow the trajectory indicated by the upper curve in the figure. A thinning at age 10 (height 31.5 feet) causes a jump by the removal of basal area, and the stand follows a new trajectory from there. Another thinning at age 15 is also shown.

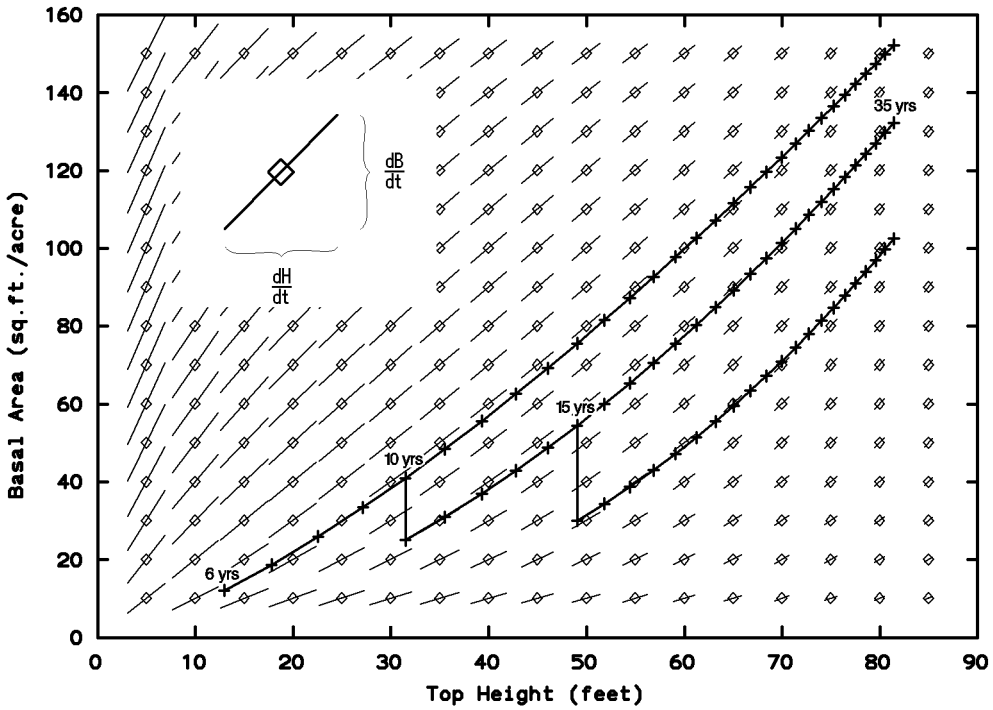

Figure 1: Vector field for Clutter's model

Any set of variables in a one-to-one relationship with the state variables could also be used as a state description. It would be essentially a different labeling for the points in the state space. For example, using (7) we could substitute V for B and obtain a model with a state vector (H, V) (doing away with the need for an output function at the same time.) This is what is behind the problem of "compatibility" addressed by Clutter (1963). Clutter used extensively the relationships between the local and global transition functions (see also Bailey and Clutter 1974), and in Sullivan and Clutter (1972) the consistency and semigroup conditions are clearly explained.

If we are not just interested in total volume, but also in volumes of sawn timber for example, we would need to augment the state vector with the mean diameter or the stocking (number of stems per hectare, N), to reflect the effect of tree size on conversion. Thus, it is common to add to (5)–(6) a mortality equation. A sufficiently general model in many circumstances takes the form

$$\frac{dH}{dt} = f_1(H) \tag{10}$$

$$\frac{dB}{dt} = f_2(H, B, N) \tag{11}$$

$$\frac{dN}{dt} = f_3(H, B, N)\ . \tag{12}$$

Many outputs can be satisfactorily estimated given H, B, and N, including volumes of various products and parameters of size distributions. The interpretation is analogous to that given above for the two-dimensional case: at any point in time a stand is characterized by its state vector $\mathbf{x} = (H, B, N)$, and the transition equations (10)–(12) indicate in what direction and how fast it will move in the three-dimensional state space; integration determines the trajectories followed in-between thinnings. Beekhuis (1966) may have been one of the first in developing a model essentially along these lines.

It has been observed that in three dimensions the trajectories tend to be concentrated near a surface (Decourt 1974, García 1988), so that the gains from the additional variable in (11) compared to (6) are not very large, although still worthwhile when covering a wide range of silvicultural regimes.

It is possible to add more state variables, although a situation of diminishing returns is soon reached. After thinning or pruning it may take some time for the trees to fully occupy the additional space that has been made available to them. It might be expected then, that immediately after the intervention a stand would grow less than another with the same H, B, and N, but not recently treated. To account for this a fourth variable, representing relative site occupancy or canopy closure, has been found useful in some instances, especially with very heavy thinning and pruning (García 1979, 1984, 1990). A model including foliar phosphorous levels as a state variable has also been developed for nutrient deficient stands (García 1988, 1989).

HEIGHT GROWTH AND SITE INDEX

A stand top height is defined in some way to represent the height of the largest trees, and in practice it may be assumed that it is little affected (within limits) by the manipulation of stand density through thinnings and initial spacing, or by pruning. I shall ignore fine points of definition (Fries 1974, Matérn 1976, Rennolls 1978), and problems with extreme treatments. Being independent of silviculture and easily measured, top height is commonly used for assessing the potential productivity (*site quality*) of forest lands. Another consequence is that the top height growth component (10) of a model constitutes a self-contained subsystem and can be developed separately.

Starting from an initial point (e. g. $H = 0$ at $t = 0$), (10) generates a unique curve for each site. To assess site quality, the *site index* is defined as the stand top height (height, for short) reached at a specified *key age* such as 20 or 50 years.

In reality, observed heights vary due to randomness in growth conditions and to measurement errors. More precisely then, since site quality is a property of the site, the site index must be defined as some average, most likely, or predicted height that a hypothetical stand growing in that site would attain, and *not* as the height that happens to be observed at the key age in a particular stand.

To estimate the site index when heights at other than the key age are available, site index curves (or equations) are used, which are a one-parameter family of curves representing the H–t trajectories for each value of the site index (Figure 2.)

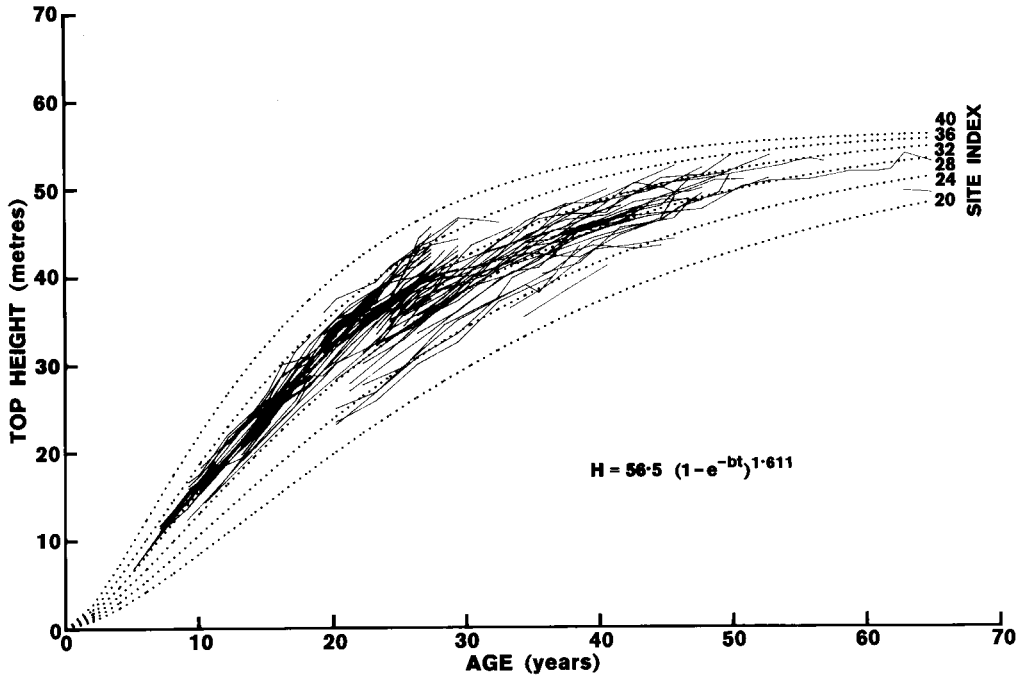

Figure 2: Site index curves (from García 1983)

The most popular site index (and height growth) equations are those of Richards [6],

$$H = a(1 - e^{-bt})^{1/c} \,, \tag{13}$$

and of Korf,

$$logH = a - (\frac{b}{t})^c \,, \tag{14}$$

this last one mostly with $c = 1$ (Schumacher [7] .) The site index concept requires one and only one parameter (a, b, c, or some function of these) to vary with site, with the rest being the same for all sites. A common choice is to have a in (13) or (14) as the indexing parameter. Then the site index curves are called *anamorphic*, because the individual curves differ only by a scaling of the H-axis. All other cases have been called *polymorphic*, which seems odd when b is the indexing parameter and the curves differ by scaling of the t-axis.

Following the state-space approach, height growth and site indexing can be modelled with a differential (or difference) equation (10), for example the differential forms of (13) or (14), containing one parameter that varies with site quality. Site index curves are obtained by integration with a suitable starting point, e. g. $H = 0$ at $t = 0$. These curves can

[6]Known also as Chapman-Richards, following Pienaar and Turnbull (1973). Used on animals by von Bertalanffy (1949, for example), and studied in detail by Richards (1959). The often-quoted generalization of Richards is to allow a negative c, and is rarely useful. The misquoting of the title of Chapman (1961) by Pienaar and Turnbull seems to have been perpetuated.

[7]The differential equation of Schumacher (1939) contains t on the right-hand-side. This can always be substituted in terms of H to obtain an equation of the form (10)

be used in the traditional way, but there is also added flexibility that can be exploited. Some possibilities are discussed in the following paragraphs.

In general, stands do not follow exactly an established site index curve. When a new height-age observation becomes available that deviates from the projected trend, the usual procedure would let the stand follow the new site index curve to which it has jumped (Figure 2). This may be justified if it is felt that the new information warrants an update of the site quality estimate. On the other hand, if it seems that an estimate based on other data is reliable, and that the deviation is due to anomalous weather or other reasons, it may be preferable to integrate from the new initial condition with the old site parameter. Thus, model projections can cut across site index curves. Obviously, then the stand will not reach the site index height at the key age, but this is how it should be.

A better fit to data might be obtained if the curves are not forced to go through the origin, and the starting point is left as another parameter to be estimated (most height-age curves are not useful for very young stands anyway.) More generally, weed competition, soil cultivation and other establishment techniques can cause substantial variation in the early height growth. These effects can be modelled by shifting the origin of the curves. With unknown establishment conditions and a floating origin, site quality could be estimated from two or more height-age measurements. The site is represented by a transition function parameter that influences the growth rate, but does not necessarily determine the size reached at a given age. We could even use a time-varying site parameter to track the effects of weather change.

Finally, it is important to realize that site index is good only for discriminating among production levels under reasonably similar growing conditions. Different conditions cause not only differences in level, but also in growth patterns, i.e., curve shapes and relationships between state variables. Stratification into moderately homogeneous growth modelling regions is often necessary, with the site index reflecting finer local distinctions within regions. In principle, an alternative would be the use of more than one site parameter, a sort of multidimensional site index.

I have used extensively a model based on (13). It happens that the local transition function can be written as a linear differential equation for a power transformation of H, a fact that is mathematically convenient:

$$\frac{dH^c}{dt} = b(a^c - H^c) \; . \tag{15}$$

In order to devise reasonable parameter estimation procedures it is necessary to have at least some crude representation of how observations deviate from the model. It was assumed that the growth rate is perturbed by environmental fluctuations through a white noise term added to the right-hand-side of (15), and that there are also measurement errors (see García (1983) or Seber and Wild (1989, Section 7.5.3) for details.) This can be seen as a transition function (3) for H where the input $\mathbf{u}$ is a white noise process, and where an output function (4) subject to error gives the measured heights in terms of the real ones. The parameters to be estimated are a, b, c, the origin, and the "environmental" and "measurement" variances. Any of these (or combinations obtained by reparametrization) can be site-dependent, i.e. specific to individual stands, or can be the same across the

entire database. A special procedure was developed that estimates all the parameters simultaneously by maximum likelihood. The computer program is freely available, and has been used successfully by a number of researchers in several countries.

PARAMETER ESTIMATION

Once a model containing unknown parameters is formulated, it is necessary to estimate those parameters. As already mentioned, a rational approach to the estimation problem requires some "metamodel" [8], a model describing the structure of the discrepancies between model and observations, usually in probabilistic language. The most complete and consistent approach to estimation, the Bayesian/decision-theoretic one, requires also to know the cost of all possible errors (a *loss function*), and a probabilistic representation of all the relevant previous information and beliefs (a *prior distribution*.) Unfortunately, the costs and consequences of errors arising from a model will vary with each use and with each user, and very rarely can be formulated in detail. The use of subjective information and beliefs is also a problem, especially when a model will be used by different people.

The more often-used methods of classical statistics do away with losses and priors, and try to find some "objective" way of making estimates based only on the (meta)model and the observations. It is not surprising that, giving the ill-defined nature of the problem stripped off its decision-making aspects, the results are not always too convincing. Estimators are built and evaluated in terms of satisfying a number of more or less ad-hock and incompatible criteria such as unbiasedness, invariance, minimum variance, minimum squared error. For a good review of the "whats" and "whys" of the various approaches to probability and statistics see Barnett (1973).

Another problem with statistics is the conditioning on models and "true" parameter values. Typical statements run like "if such and such a model is true, then...". Given that, by definition, there is no such thing as a "true model", the conclusions must be treated with some skepticism. Having said all this, I must emphasize that statistical theory and methods are extremely useful, and even indispensable in growth modelling. Used with care and understanding, statistical models and techniques serve to formulate and approach problems of incomplete information in a rational way, and to produce usually reasonable, if not perfect solutions.

Perhaps the most generally applicable and useful estimation method is the method of maximum-likelihood (ML). The likelihood function is the probability of generating the observed data with the (probabilistic) model, viewed as a function of the unknown parameters. The ML estimate is that value of the parameters for which the likelihood function attains its maximum value. Intuitively, of all possible values of the parameters, the ML estimates are those having the largest probability of giving rise to a sample close to the actually observed one (Bard 1974).

Under some conditions ML estimators are shown to possess a number of "good" properties, mainly asymptotically, for large samples. In certain situations they approximate or

[8] In recent simulation literature this term is being used with a different meaning, a simplified model for the behavior of a more complicated model

coincide with Bayesian results and with other estimation methods such as least-squares, and generally produce acceptable estimates in most circumstances. But the usefulness of the ML method in modelling is largely due to two characteristics. First, no matter how complicated the model, there is always a well-defined procedure with which parameter estimates can be obtained, at least in principle. One "simply" establishes the likelihood function and finds its maximum. Second, the ML method is invariant under re-parametrization, and any quantities computed from the model with the ML estimates substituted for the parameters are ML estimates for those quantities.

Least-squares is a special case of ML when the residuals for which the sum of squares is minimized are independent, and all have the same normal distribution. In linear regression, least-squares can be justified also as giving the smallest variance among all the unbiased estimators that are linear functions of the observations, provided that the residuals are independent and all have mean zero and the same variance, even if they are not normal (the Gauss-Markov theorem.)

Bard (1974) discusses estimation in dynamic systems (see also Seber and Wild 1989, Chapter 7.) Consider a model

$$\frac{d\mathbf{x}}{dt} = \mathbf{f}(\mathbf{x}, \theta) \; , \tag{16}$$

where the state $\mathbf{x}$ is observable and θ is a vector of parameters to be estimated. The data consists of values of $\mathbf{x}$ at various t, but we do not usually directly measure the values of the derivatives. A first approach is to calculate approximate values of the derivatives in (16) by differencing adjacent data values. If $(\mathbf{x}_1, t_1)$ and $(\mathbf{x}_2, t_2)$ are two consecutive data points, then $(\mathbf{x}_2 - \mathbf{x}_1)/(t_2 - t_1)$ is an approximation to $d\mathbf{x}/dt$ at $t = (t_1 + t_2)/2$. Where a treatment (e. g. thinning) has caused a change of state between two measurements, that pair of measurements should not be used. Using all the relevant consecutive measurement pairs, the equation in (16) for each of the state variables can then be fitted by least-squares, for example.

Fitting the equations separately may not give the best results if their errors are correlated, and it is not feasible at all if some parameter appears in more than one equation. For these *multiresponse* situations, using Bayesian arguments Box and Draper (1965) recommended estimation by minimizing the determinant of the matrix of sample joint moments (see also Hunter (1967), Bates and Watts (1985), Stewart (1987), Kang and Bates (1990), and references therein.) Bard (1974, Section 4-9) obtained the same determinant criterion with the ML method. Note also that measurement errors would cause autocorrelation, see the discussion following (19), below.

Because of its simplicity, the direct fitting of (16) with approximate derivatives may be the best option in many instances, although Bard objects that its accuracy is severely limited, and its errors difficult to assess. If the model uses difference instead of differential equations and the spacing of the observations corresponds to the difference time-intervals, then there is no approximation error. With frequent observations uniformly spaced, the errors may be acceptable. However, the method cannot be used effectively if the separation between data points is large and/or irregular.

The second approach is based on integrating (16), obtaining the global transition function

$$\mathbf{x} = \mathbf{F}(t - t_0, \mathbf{x}_0, \theta) \; . \tag{17}$$

(The third approach discussed by Bard, numeric integration of data, does not seem applicable to our problems.) The initial conditions $\mathbf{x}_0$ at t_0 are usually known (measured); if not, they would be included in the problem as additional parameters to be estimated. Only instances where (17) can be obtained analytically in closed form seem to have been considered in forestry, although, in principle, numerical integration could also be used. Often, the model is given from the start as functions of time like these, in site index applications and in growth models of the static type.

A characteristic of these problems is that the data consists of many disjoint series of observations, each of them modelled by (17), possibly with different initial conditions. This is sometimes called *panel* data, which is a combination of *cross-sectional* data, where single observations are made on different individuals, and *longitudinal* data, where a sequence of observations are made on a single individual. As usual, obtaining good estimators for the unknown parameters requires reasonable assumptions about the error structure. Specifically, to apply ML or Bayesian methods we need an approximation to the joint probability density of all the observations (the likelihood function.) For least-squares we need to form independent and identically distributed residuals.

Usually correlation across individuals is not too important, and can be ignored, so that only the distribution of the sequence of observations within one individual needs to be considered (but see García 1983.) We observe $\{(t_1, \mathbf{x}_1), (t_2, \mathbf{x}_2), \ldots, (t_m, \mathbf{x}_m)\}$. This would follow (17), except for two main sources of variation: measurement (possibly including sampling) errors, and an unpredictably varying environment.

A common approach is to use least-squares under the model

$$\mathbf{x}_i = \mathbf{F}(t_i - t_0, \mathbf{x}_0, \theta) + \epsilon_i \ \ ; \quad i = 1, 2, \ldots, m \ . \tag{18}$$

This has been done in univariate models, or one equation at a time, where the $\mathbf{x}_i$ are scalars. Minimization of the determinant criterion could be used in multiresponse situations.

This method can be expected to work well if the ϵ_i are approximately independent and identically distributed. This is the case if the deviations from the model are predominantly due to measurement error (transformations can be used to equalize the variances.) On the other hand, if the trajectories are perturbed by environmental variation these perturbations have a cumulative effect, causing the ϵ_i to be correlated, and to have variances increasing with time. To see why, observe that the realized value of $\mathbf{x}$ at time t_{i+1} depends of the value at t_i, through (17), and of the perturbations between t_i and t_{i+1}, while in turn , the value at t_i deviates from the nominal trajectory due to perturbations occurred before t_i.

Failing the assumption of uncorrelated errors, the least-squares estimators based on (18) can be inefficient, and errors can be grossly underestimated (Sullivan and Reynolds 1976, Seber and Wild 1989, Section 7.4). One alternative is to allow for correlated errors with techniques such as generalized least-squares or autoregresive time series modelling (Seber and Wild 1989). This may not be very effective because of the introduction of large numbers of unknown correlation parameters and/or arbitrary error structures. At best, as Seber and Wild (1989, p.344) puts it, it is rather clumsy, and a more reasoned approach may be desirable.

Instead of (18), we may consider the error structure

$$\mathbf{x}_i = \mathbf{F}(t_i - t_{i-1}, \mathbf{x}_{i-1}, \theta) + \epsilon_i \ ; \quad i = 1, 2, \ldots, m \ . \tag{19}$$

It is reasonable to assume that environmental fluctuations are of the white noise type (but, again, see García 1983.) That is, fluctuations within some time interval are independent from those in other non-overlapping intervals, at least approximately and on a suitable time scale (years.) If, in addition, measurement errors are comparatively unimportant the ϵ_i in (19) can be taken as uncorrelated.

Under the preceding assumptions, if the observations are uniformly spaced in time then we have essentially the finite differences situation already discussed, and least-squares or minimization of the determinant of the moment matrix for (19) may be appropriate. If, on the contrary, the separation between observations varies widely, it may be worthwhile to model how the variance of ϵ_i changes with the length of the interval. This has been done, for example, by García (1979, 1984) integrating differential equations that include a white noise process, and by Candy (1989) using generalized linear models. Several approaches are discussed by Seber and Wild (1989, Section 7.5), and by Jones and Ackerson (1990).

When both measurement and environmental sources of variation are important the situation is more complicated. The ϵ_i in (19) become correlated, because consecutive intervals share a common measurement. García (1983) modelled the two error causes, although the ML method rarely has been able to estimate separately the two variances.

In general, as indicated above when discussing height growth, the error structure for a dynamic system can be modelled through an input $\mathbf{u}$ in the transition function (3) where $\mathbf{u}$ is a stochastic process representing environmental fluctuations, and an output function (4) for the observations that includes measurement errors. The problem resides in devising models like this that are sufficiently realistic, and at the same time mathematically tractable.

In trying to improve estimation methods the mathematics becomes increasingly involved, and computation effort rises rapidly. It is difficult to know in a particular instance how much refinement will produce worthwhile improvements. It may be argued, however, that with costly and time-consuming field work and increasing computing power, there are few excuses for not attempting to squeeze as much information out of the data as we possibly can.

THE MULTIVARIATE RICHARDS MODEL

The state-space theory explained above is completely general, and there are many possibilities in choosing state vectors and equations. The following example may serve as an interesting case study.

When initiating the development of a series of regional growth models for radiata pine plantations in New Zealand, we were confronted with the lack of any satisfactory theory that would dictate a particular form for the transition equations. In addition, there was sufficient data, representative of a wide range of initial spacings and thinning regimes. It

was decided then to use empirical models flexible enough to fit the observed development patterns, unconstrained by any dubious preconceived ideas. Stand-level state variables based on top height, basal area, and stems per hectare, possibly adding green crown level or other measure of canopy closure, were thought adequate, at least in a first stage. To obtain efficient estimation procedures it was desirable also for the form of the model to be mathematically convenient. In particular, it should be possible to integrate the transition equations analytically.

Linear differential equations (DEs) are very convenient and easily handled, but not very flexible (García 1974). A one-dimensional DE is called the Mistcherlich or monomolecular model, and produces curves without an inflection point that rarely represent well tree growth. Somewhat better results were obtained with a system of linear DEs in the logarithms of the state variables. This is a generalization to several variables of the Gompertz model, and Minowa (1982, 1983a, 1983b) developed it further.

The final model used DEs in power transformations of the form

$$H^{\alpha}B^{\beta}N^{\gamma} \ . \tag{20}$$

This can be seen as an n-dimensional generalization of the Richards model if we define [9]

$$\mathbf{x}^{\mathrm{C}} \equiv \exp[\mathrm{C} \ln \mathbf{x}] \ . \tag{21}$$

This is a vector of transformations like (20) for the n state variables in $\mathbf{x}$, with the power coefficients collected in the $n \times n$ matrix C. The model can then be written as

$$\frac{d\mathbf{x}^{\mathrm{C}}}{dt} = \mathrm{B}(\mathbf{a}^{\mathrm{C}} - \mathbf{x}^{\mathrm{C}}) \ , \tag{22}$$

which is analogous to (15) (García 1979). The coefficients in the vector and matrices $\mathbf{a}$, B and C may be functions of the site index, containing parameters to be estimated.

A nice side-effect of (20) is that many variables used in forestry are proportional to some instance of this. For example, the mean dbh (diameter at breast height, $H^0B^{0.5}N^{-0.5}$), average spacing ($H^0B^0N^{-0.5}$), the Hart-Becking relative spacing ($H^{-1}B^0N^{-0.5}$), Reineke's density index ($H^0B^{0.8}N^1$), some approximations to the volume ($H^1B^1N^0$ plus a constant, or $H^{\alpha}B^{\beta}N^0$ in logarithmic volume equations.) Using any independent set of these for $\mathbf{x}$ does not make any difference to the form of (22), saving us having to make some ackward decisions. I have used $\mathbf{x} = (H, B, N)$ or $\mathbf{x} = (H, B, N, R)$, where R is some measure of relative canopy closure (García 1984, 1988).

Integration of (22) between t_0 and t, with the initial condition $\mathbf{x}(t_0) = \mathbf{x}$, gives the global transition function

$$\mathbf{x} = [\mathbf{a}^{\mathrm{C}} + e^{\mathrm{B}(t-t_0)}(\mathbf{a}^{\mathrm{C}} - \mathbf{x}^{\mathrm{C}})]^{\mathrm{C}^{-1}} \tag{23}$$

that predicts $\mathbf{x}$ for any time t such that there are no changes of state caused by thinning or pruning between t_0 and t. The matrix exponential has been computed as

$$e^{\mathrm{B}(t-t_0)} = \mathrm{P}^{-1}e^{\Lambda(t-t_0)}\mathrm{P}$$

with the eigenvalue decomposition $\mathrm{B} = \mathrm{P}^{-1}\Lambda\mathrm{P}$, where Λ is diagonal (other alternatives are reviewed by Moler and Loan 1978.)

[9] The logarithm of a vector is the vector of logarithms of its elements, and similarly for exp.

(22) constitutes a fairly general and flexible class of models that could be used for many dynamic systems. A number of details and specializations are relevant in a specific growth modelling application. The selection of state variables has already been mentioned. According to (10), the height equation should not involve other state variables. This is achieved by having only one non-zero coefficient in the first rows of B and C, if H is the first component of $\mathbf{x}$. This first equation is then the same as (15), and as the height growth / site index sub-system it is estimated first, separately from the rest of the model. Other coefficients may also be fixed at zero or other values, for example in order to restrict N to be non-increasing (García 1984).

It is also necessary to decide how the site index will enter in the equations. For the North Island Sands region the best height growth model was of the anamorphic type, that is, in (13)–(15) the asymptote a is the indexing parameter that varies with site quality. It seemed natural to take the other asymptotes in $\mathbf{a}$ of (22) as proportional to a, and B and C as independent of site. Analysis of residuals showed this assumption to be satisfactory. In the other 6 regions for which models were developed the site indexing parameter was b, so that site acts by altering uniformly the height growth rates or, equivalently, by changing the time scale. Good results were obtained doing the same with the other rates by keeping $\mathbf{a}$ and C constant and making B proportional to b (or, more conveniently, by scaling t according to b.) An exception was the Hawke Bay region, where a bias was detected and multipliers b and b^k had to be used on different equations.

Some extensions of the basic model have also been used. These involve adding functions of state variables or other quantities as multipliers applied to some of the equations in (22), in such a way that the system can still be integrated analytically in closed form. Thus, a better-behaved model was obtained by introducing the R variable through a multiplier, instead of including it in (21). A similar approach was used to model nutrient deficiency with a state variable representing the foliar concentration of phosphorous. For details, see García (1989). More recently, the effects of using various kinds of genetically improved seed have been modelled with multipliers.

The method of García (1983) discussed in the previous section has been used to estimate the parameters in the height growth sub-models. These parameters are considered as known when estimating the remaining ones in (22)–(23). This last estimation problem is easier in that the large number of site-dependent parameters is no longer present, although being multivariate makes it more complicated in other respects.

One possibility for estimation is to use derivative approximations in (22). This was done for radiata pine in Chile in a thesis by de la Maza (1979), under my supervision. The model was slightly different in that the derivatives were with respect to top height, instead of time. This was to get around a problem of uncertain measurement dates, and at the same time to get rid of site (assuming a time-scaling effect, following Beekhuis 1966.) Given tentative values for C, the other parameters of (22) were obtained by linear regression for each equation. With these, various C were evaluated with the likelihood function corresponding to normal independently distributed errors on the right-hand-side.

A more elaborate procedure estimates all the parameters simultaneously (including C), avoids the inaccuracy of derivative approximations, and extracts the maximum of information from data with arbitrary intervals between measurements (García 1979, 1984). It is analogous to that for height growth in García (1983), but it applies to the multivariate

situation, it does not model observation errors in addition to the environmental variation, and it does not need to cope with the large number of site-indexing parameters. Random perturbation terms are added to (22), and the resulting *stochastic differential equations* are integrated to obtain the full likelihood function in closed form. The maximum of the likelihood function is found with a general-purpose optimization routine. The method has performed well, and the relative complexity of its development was justifiable by the need to produce a series of several growth models. An interesting refinement has been the use of computer-generated derivative routines instead of the finite-difference approximations for the likelihood derivatives previously used in the optimization (García 1991).

The performance of the models has been very satisfactory. A model for Douglas Fir was also produced with the same methodology. They form part of STANDPAK (Gordon 1992) and other software packages, and together with EARLY (West et al. 1982, commonly used for silviculture scheduling in young stands) form the basis for most of New Zealand plantation forestry planning. Despite having been chosen primarily for flexibility and mathematical convenience, the multivariate Richards model, like its univariate counterpart, generally behaves in a biologically reasonable way under extrapolation and in limiting situations.

PATTERNS OF GROWTH

Analysis of and experience with the New Zealand growth models and the data used in them has provided some insights on the development of forest stands. These observations may be useful for improving the accuracy of future models, and for modelling in data-poor situations, although we are still far from having a complete and fully satisfactory theoretical modelling framework. Only a brief synthesis is presented here, for more details see García (1990). We deal with good sites, where light is likely to be the main limiting factor; different patterns may occur where growth is limited by the availability of water or nutrients.

It was found that in closed stands the gross volume increment (after adjusting for mortality, if any) is nearly independent of age, at least within practical rotation lengths. Only a relatively small, roughly linear decrease with increasing spacing has been detected. Except for some curvature at high densities caused by mortality, after stand closure the volume–age trends are largely linear. This goes against the familiar textbook vision of sigmoidal curves with well-defined asymptotes, although it agrees with the common observation that, within limits, stand density has little or no effect on total gross volume production. In retrospect, the pattern is also evident when reviewing some European thinning experiments. The constant gross production would be expected from the full site occupancy and the constant input of light and other resources per unit area, provided that tree efficiency does not change too much with size. A plausible explanation for the reduced growth at wider spacings might be losses and structural and maintenance overheads associated with longer roots and branches.

It is necessary also to predict growth for stands that are not fully closed, as is the case at young ages, and immediately following heavy thinning and pruning. A state variable depicting relative closure may be used for a simple mechanistic model of stand behavior.

The *closure* variable represents the amount of assimilating material (foliage, roots) relative to that in a closed stand. It is assumed that closure starts at a small value proportional to the number of trees per hectare, and gradually increases up to unity when the stand closes, following some transition function. Thinning reduces the current closure in proportion to the relative basal area or volume removed, and pruning also reduces closure in some specified way. After thinning or pruning, closure increases again toward one, according to the transition function. The stand's volume increment is determined by its *occupancy*, defined as the increment relative to the (gross) increment in a fully closed stand. The occupancy is linked to the closure by a non-linear function, since it is well-known that moderate changes in closure have a small effect on growth. Together with the increment-spacing relationship for closed stands discussed above, this suffices as a somewhat crude but not too unrealistic model for the dynamics of thinned and pruned stands. Natural mortality is unimportant in intensively managed stands, but it would need to be included to extend the model applicability beyond these.

The model just outlined describes the state with a "closure" variable, in addition to the usual stem-based characteristics such as height, volume and spacing. In practice, models based only on stem variables often perform very well, due to an interdependence of variables analogous to multicollinearity in linear regression. The concentration of data points along a surface in a three-dimensional state space noticed by Decourt (1974) has already been mentioned. García (1990) shows how a normally close relationship between closure and volume in the data can make the current volume to appear as a good explanatory variable for predicting volume increment. However, this one-dimensional transition function for volume, the so-called Langsaeter hypothesis, can fail for "atypical" treatments.

Another subject into which the state space approach can throw some light is that of stand density indices and self-thinning. It is often assumed that competition-induced mortality causes stands to approach a limiting linear relationship between $\log N$ and the logarithm of some other variable. Distance to the limit line is sometimes used as an index of stand density. There are three commonly-used hypothesis of this type. Reineke's (1939) density index is based on

$$\log N + \alpha \log D = \text{constant} \ , \tag{24}$$

where D is the mean dbh and the constant α is usually about 1.6. The relative spacing or Hart-Becking index, i. e. the average spacing divided by top height, is often assumed to tend to a limiting value (Beekhuis 1966), implying

$$\log N + 2 \log H = \text{constant} \ . \tag{25}$$

Perhaps the most famous and most controversial of these hypotheses is the "3/2 Law of Self-Thinning"(Hutchings 1983, Weller 1987, Lonsdale 1990), which says that stands of plants reach a limiting density such that the average plant mass $w \equiv W/N$ is proportional to $N^{-3/2}$, i.e.,

$$\log N + \frac{2}{3} \log w = \text{constant} \ . \tag{26}$$

In forestry, stem volume is commonly substituted for the mass [10].

[10]Weller (1987) suggests that the division by N and the logarithmic transformations may have something to do with the relationships commonly observed. A good exercise is to generate random samples of W and N independently out of some distribution(s), plot a scattergram of (W, N) pairs, and then plot $\log(W/N)$ over $\log N$.

Consider a three-dimensional state space with coordinates H, D and N, and assume that, as is generally the case, this is an acceptable state description (similar arguments would apply in more general spaces.) Stands starting with the same initial stocking N_0, that is from a point $(N_0, 0, 0)$ or $(N_0, 0, 1.4)$ for example, would follow the same trajectory if left unthinned. The projection of this trajectory onto the (D, N) plane gives a relationship between these two variables. It is plausible that on a logarithmic scale this relationship eventually becomes roughly linear, approaching Reineke's limiting line. Similarly, the projection onto the (H, N) plane determines the course of the relative spacing, with a hypothetical limit $2 \log H + \log N = \text{constant}$. Hence, the limiting lines of these indices are projections onto certain planes of a linear approximation to the trajectory, for large t, in the equivalent transformed space of $(\log H, \log D, \log N)$. An infinity of other limiting lines and density indices could be obtained by projection onto specific planes, or by linear transformations of the logarithmic coordinates. For example, using $D^2 H$ as a proxy for w, the self-thinning line is a projection onto the plane $(2 \log D + \log H, \log N)$ or $\log D - 2 \log H = 0$. Incidentally, with the customary constants of $3/2$ for the self-thinning line and about 1.6 for Reineke's, the three lines are incompatible, that is, there cannot be a trajectory satisfying exactly the three relationships. An acceptable approximation is possible, though.

Therefore, for natural, unmanaged stands, where most of the self-thinning and stand density work has been done, acceptable approximations to the limiting lines can be expected, provided that the initial stocking does not vary too much. The three-dimensional trajectories are bunched together, and projections onto any two-dimensional plane will give reasonable relationships between pairs of variables.

This is not the case under wide variations in stocking. The stands start then from points $(N_0, 0, 0)$ or $(N_0, 0, 1.4)$ for a range of values of N_0, which are scattered over a line in the (H, D, N) space. For a given time interval, the transition function (assumed continuous) transforms this line into a curve, the curve into another curve, etc., generating a surface embedded in the three-dimensional state space. Thus, the trajectories originating from the initial line generate a surface and not a spatial curve, so that most projections will not work well except in very special situations. It has been observed that thinning moves points roughly along the surface, without deviating to far from it, with the result that the distribution of measurements from thinned stands is also fairly flat (the "Decourt effect".) Briegleb (1952) recognized that for managed stands, relating limiting (or "standard") N to D or to H alone was inadequate, and looked at the relationships between standard N and both D and H together.

It is natural to assume that, at least as a first approximation, limiting lines for the trajectories in $(\log H, \log D, \log N)$ space generate a limiting plane for the state surface,

$$\log N + a \log D + b \log H = c \, . \tag{27}$$

This expression would serve as a density index like those suggested by Briegleb, and projections onto planes perpendicular to it, for example onto $(\log D^{(a/b)} H, \log N)$, would yield lines. If the individual limiting trajectory lines embedded in the plane were compatible with both (24) and (25), it is found that the coefficients in (27) would have to satisfy

$$\frac{2a}{2 - b} = \alpha \, . \tag{28}$$

Equation (27) was fitted by least-squares to 20 values from Briegleb's Table 2, giving $a = 2.76$ and $b = -1.47$. The fitted values agree to within 2 stems per acre with all the values in Briegleb's table, and (28) gives $\alpha = 1.59$, remarkably close to Reineke's 1.6.

Later, a preliminary analysis of some radiata pine data gave $a = 3.28$ and $b = -2.29$ ($\alpha = 1.53$), estimated through principal component analysis (PCA) of the log-transformed variables. PCA is preferred to regression because the three variables are subject to error; see for example Weller (1987), who used it on two variables for self-thinning lines. One eigenvalue was very small, indicating that the data was close to the plane (27). When only unthinned stands were included, two of the eigenvalues were small, suggesting clustering around a line, as predicted by the theory. Two-variable PCA of the unthinned data gave an estimate of 1.58 for α in (24), but only 1.72 instead of 2 for the coefficient of $\log H$ in (25). This work done in 1976-77 was not continued and published, among other reasons, because of the paucity of unthinned stands in our database, and because self-thinning was only of academic interest under New Zealand conditions. Some of the ideas, however, were taken up and explained by Minowa (1983).

The specific slopes of the self-thinning and limiting density lines (or planes), if real, remain intriguing empirical facts. The theories that have been advanced in the literature to explain them are not entirely convincing. But the existence of two-dimensional projections for any individual trajectory or its asymptote, and in some instances for groups of trajectories that pass through some fixed spatial curve, are simply properties of dynamic systems.

DISCUSSION, SUMMARY AND CONCLUSIONS

Explained and understood, the state-space ideas are very simple, and likely to be seen as obvious or well-known. Much confusion still exists, however, about the proper handling of time functions. Ignoring the need for appropriate multi-dimensional state descriptions may be behind "... the biologist's tendency to put a lot of related quantities together in some arbitrary way to form the Coefficient of This and the Index of That..." (S. C. Pearce in the discussion of Mead and Riley 1981.) The System Theory point of view may help to understand better than in the past the foundations of growth modelling and the various approaches that have been used.

The generalized abandonment of stand-level modelling in favor of the more fashionable individual-tree approaches may not be justified, at least not for homogeneous even-aged pure stands. Large state vectors are likely to contain mostly redundant information, with consequent losses in precision (overparametrization), besides being costly or impossible to obtain their values accurately from routine field sampling. In addition, ignoring microsite correlations in individual-tree models can produce unrealistic and misleading results. Improved methods can produce good predictions for the dynamics of stand-level quantities, and it seems prudent to transfer the more questionable distribution estimates to the output function.

Good site index curves can be obtained, making full use of measurements taken at arbitrary ages, with simultaneous ML estimation of all the parameters. There is some scope

for more sophisticated site-indexing approaches using floating origins or multivariate indices.

There are some conceptual problems in Statistics of which users should be aware in order to understand the properties and limitations of statistical methods, and to interpret results correctly. Growth data presents unique statistical characteristics that hamper the development of fully satisfactory estimation methods. Yet, at least some rough statistical model for the error structure is required to choose estimates in a rational way. Ad-hock procedures should be avoided; the implied assumptions are likely to be unsound.

The principles described formed the basis for a major growth modelling project in New Zealand, producing 10 growth models for radiata pine and one for Douglas Fir. The methodology proved to be feasible and effective, and provided a consistent framework for the development of the series of models. Excellent predictions have been obtained over a wide range of treatments and growing conditions.

Gross volume increment in closed stands has been found to be independent of age for practical rotation lengths, decreasing somewhat with tree spacing. A simple mechanistic model is obtained by adding variables representing closure and occupancy for stands that are not fully closed. These ideas may be used for prediction in data-poor situations, and may contribute to better growth models in the future.

It was shown how within a state-space framework it is possible to throw some light onto the intriguing subject of self-thinning and stand density. Some observations are just consequences of general properties of dynamic systems. For managed stands, a "self-thinning plane" would be an improvement over the various self-thinning or limiting lines. Many mysteries disappear when we step-out of Flatland (Peterson 1988, p.82–84).

REFERENCES

Bailey, R. L. and Clutter, J. L. (1974). Base-age invariant polymorphic site curves. *Forest Science*, *20*, 155–159.

Bard, Y. (1974). *Nonlinear Parameter Estimation*. New York: Academic Press.

Barnett, V. (1973). *Comparative Statistical Inference*. London: Wiley.

Bates, D. M. and Watts, D. G. (1985). Multiresponse estimation with special application to linear systems of differential equations (with Discussion). *Technometrics*, *27*, 329–353.

Beekhuis, J. (1966). Prediction of yield and increment in *pinus radiata* stands in New Zealand. Technical Paper 49, Forest Research Institute, NZ Forest Service.

Box, G. E. P. and Draper, N. R. (1965). The Bayesian estimation of common parameters from several responses. *Biometrika*, *52*, 355–365.

Briegleb, P. A. (1952). An approach to density measurement in Douglas-Fir. *Journal of Forestry*, *50*, 529–536.

Bunnell, F. L. (1989). Alchemy and uncertainty: What good are models? General Technical Report PNW-GTR-232, USDA Forest Service.

Candy, S. G. (1989). Growth and yield models for Pinus radiata in Tasmania. *New Zealand Journal of Forestry Science*, *19*, 112–133.

Chapman, D. G. (1961). Statistical problems in dynamics of exploited fisheries populations. In *Fourth Berkeley Symposium — Mathematical Statistics and Probability, Proceedings*, pages 153–168, Berkeley. University of California Press.

Clutter, J. L. (1963). Compatible growth and yield models for loblolly pine. *Forest Science*, *9*, 354–371.

Clutter, J. L., Fortson, J. C., Pienaar, L. V., Brister, G. H., and Bailey, R. L. (1983). *Timber Management: A Quantitative Approach*. London: Wiley.

Clutter, J. L. and Lenhart, J. D. (1968). Site index curves for old-field loblolly pine plantations in the Georgia Piedmont. Report No. 22 – Series 1, Georgia Forestry Research Council.

de la Maza, J. E. (1979). Un modelo dinámico de rodal con aplicación a raleos en pino insigne. *Ciencia Forestal*, *1*(3), 11–27.

Decourt, N. (1974). Remarque sur une relation dendrométrique inattendue — Conséquences méthodologiques pour la construction des tables de production. *Annales des Sciences Forestières*, *31*, 47–55.

Fries, J. (1974). Praktisk bonitering. *Sveriges Skogsvårdsförbunds Tidskift*, *72*, 559–563.

García, O. (1974). Sobre modelos matemáticos de rodal. Informe Técnico 48, Instituto Forestal, Chile.

García, O. (1979). Modelling stand development with stochastic differential equations. In Elliott, D. E. (Ed.), *Mensuration Systems for Forest Management Planning*, pages 315–334, Forest Research Institute Symposium No. 20. New Zealand Forest Service.

García, O. (1983). A stochastic differential equation model for the height growth of forest stands. *Biometrics*, *39*, 1059–1072.

García, O. (1984). New class of growth models for even-aged stands: *Pinus radiata* in Golden Downs Forest. *New Zealand Journal of Forestry Science*, *14*(1), 65–88.

García, O. (1988). Experience with an advanced growth modelling methodology. In Ek, A. R., Shifley, S. R., and Burke, T. E. (Eds.), *Forest Growth Modelling and Prediction*, pages 668–675. USDA Forest Service, General Technical Report NC-120.

García, O. (1989). Growth modelling — New developments. In Nagumo, H. and Konohira, Y. (Eds.), *Japan and New Zealand Symposium on Forestry Management Planning*, pages 152–158. Japan Association for Forestry Statistics.

García, O. (1990). Growth of thinned and pruned stands. In James, R. N. and Tarlton, G. L. (Eds.), *New Approaches to Spacing and Thinning in Plantation Forestry: Proceedings of a IUFRO Symposium, Rotorua, New Zealand, 10–14 April 1989*, pages 84–97. Ministry of Forestry, FRI Bulletin No. 151.

García, O. (1991). A system for the differentiation of Fortran code and an application to parameter estimation in forest growth models. In Griewank, A. and Corliss, G. F. (Eds.), *Automatic Differentiation of Algorithms: Theory, Implementation, and Application*, pages 273–285. Society for Industrial and Applied Mathematics.

García, O. (1992). What is a diameter distribution? In Minowa, M. and Tsuyuki, S. (Eds.), *Proceedings of the Symposium on Integrated Forest Management Information Systems*, pages 11–29. Japan Society of Forest Planning Press.

Gordon, A. (1992). Design and implementation of a micro-computer stand evaluation package. In Wood, G. and Turner, B. (Eds.), *Integrating Forest Information Over Space and Time, Proceedings of the International IUFRO Conference*, pages 178–187, Canberra, Australia. ANUTECH Pty Ltd.

Hunter, W. G. (1967). Estimation of unknown constants from multiresponse data. *Industrial & Engineering Chemistry. Fundamentals*, *6*, 461–463.

Hutchings, M. (1983). Ecology's law in search of a theory. *New Scientist*, *(16 June)*, 765–767.

Jones, R. H. and Ackerson, L. M. (1990). Serial correlation in unequally spaced longitudinal data. *Biometrika*, *77*, 721–731.

Kalman, R. E., Falb, P. L., and Arbib, M. A. (1969). *Topics in Mathematical System Theory*. New York: McGraw-Hill.

Kang, G. and Bates, D. M. (1990). Approximate inferences in multiresponse regression analysis. *Biometrika*, *77*, 321–331.

Lonsdale, W. M. (1990). The self-thinning rule: Dead or alive? *Ecology*, *71*, 1373–1388.

Matérn, B. (1976). Om skattning av övre höjden. *Sveriges Skogsvårdsförbunds Tidskift*, *74*, 51–53.

Mead, R. and Riley, J. (1981). A review of statistical ideas relevant to intercropping research (with Discussion). *Journal of the Royal Statistical Society, A*, *144*, 462–509.

Minowa, M. (1982). A theoretical approach to forest growth modeling — (I) The log-Mistcherlich theory [in Japanese]. *Journal of the Japanese Forestry Society*, *64*, 461–467.

Minowa, M. (1983). A theoretical approach to forest growth modeling — (II) Further discussion on the self-thinning model [in Japanese]. *Journal of the Japanese Forestry Society*, *65*, 135–142.

Moler, C. and Loan, C. V. (1978). Nineteen dubious ways to compute the exponential of a matrix. *SIAM Review*, *20*, 801–836.

Padulo, L. and Arbib, M. A. (1974). *System Theory*. Washington, D.C.: Hemisphere Pub. Co.

Peterson, I. (1988). *The Mathematical Tourist: Snapshots of Modern Mathematics*. New York: Freeman.

Pienaar, L. V. and Turnbull, K. J. (1973). The Chapman-Richards generalization of von Bertalanffy's growth model for basal area growth and yield in even-aged stands. *Forest Science*, *19*, 2–22.

Rennolls, K. (1978). "top height"; its definition and estimation. *Commonwealth Forestry Review*, *57*, 215–219.

Richards, F. J. (1959). A flexible growth function for empirical use. *Journal of Experimental Botany*, *10*, 290–300.

Schumacher, F. X. (1939). A new growth curve and its application to timber-yield studies. *Journal of Forestry*, *37*, 819–820.

Seber, G. A. F. and Wild, C. J. (1989). *Nonlinear Regression*. New York: Wiley.

Stewart, W. E. (1987). Multiresponse parameter estimation with a new and noninformative prior. *Biometrika*, *74*, 557–562.

Sullivan, A. D. and Clutter, J. L. (1972). A simultaneous growth and yield model for loblolly pine. *Forest Science*, *18*, 76–86.

Sullivan, A. D. and Reynolds, Jr., M. R. (1976). Regression problems from repeated measurements. *Forest Science*, *22*, 382–385.

von Bertalanffy, L. (1949). Problems of organic growth. *Nature*, *163*, 156–158.

Weller, D. E. (1987). A reevaluation of the -3/2 power rule of plant self-thinning. *Ecological Monographs*, *57*, 23–43.

West, G. G., Knowles, R. L., and Koehler, A. R. (1982). Model to predict the effects of pruning and early thinning on the growth of radiata pine. FRI Bulletin 5, Forest Research Institute, Rotorua, New Zealand.

Zadeh, L. A. (1969). The concepts of system, agregate, and state in System Theory. In Zadeh, L. A. and Polak, E. (Eds.), *System Theory*. New York: McGraw-Hill.