Growth models for Sitka spruce in Ireland

Lance. R. Broad[a] and Ted Lynch[b]

Abstract
Ad hoc techniques were used to fit stand-level growth models for Sitka spruce in Ireland in the presence of sampling biases. Panel data, sourced by aggregation over a set of repeated measures experimental designs, admitted sampling biases when considered as a data set for yield modelling purposes. Remedial action towards sampling imbalances was taken, whenever possible, through the use of weighting techniques. Further action to address data omissions involved using data substitution. The growth component model used was the multivariate Bertalanffy-Richards model formulated by García. The model allows aggregated state-space representations of stand development to be formulated and fitted. For stand simulations, growth equations were augmented with functions accepting state variables as arguments and compute basal area reduction during any thinning, stand volume and volume assortments.

Keywords: Assortments, growth models, multivariate growth, Sitka spruce, weighting techniques.

Introduction
Sitka spruce (*Picea sitchensis* (Bong.) Carr.) was first introduced into Ireland from British Columbia in the 1830s. Publicly funded afforestation followed in 1904 (Joyce and OCarroll 2002). The species has flourished within its Irish setting and a prosperous forest industry based on its utilisation has developed. Since the 1960s, this industry has relied heavily upon Forestry Commission (FC) yield tables and the stand management concepts surrounding them (Johnston and Bradley 1963).

However, there are aspects of modern forest management that the FC yield tables are less equipped to deal with Specifically, the FC tabular models are not well suited to introducing ideas from economics regarding the management of natural resources and optimal stand or forest management. Consequently, the aim of this work was to develop an alternative growth projection mechanism for Sitka spruce that is amenable to simulating a wide range of management alternatives specific to Irish conditions. Two Sitka spruce models were developed – a model for un-thinned stands and a model for thinned stands.

Reliably constructed growth models that permit growth and yield forecasting are crucially important for management of forest plantations. Silvicultural and economic planning at the stand and forest estate level are contingent on being able to forecast growth accurately at the stand level. Until the pioneering work of Clutter (1963) the problem of determining annual increments within forest stands, i.e. growth assessment, was conducted separately from that of determining stand aggregates or yield assessment. Clutter showed that the increments could be derived from the

[a] Technical Forestry Services, New Zealand.

[b] Coillte Research and Development, Newtownmountkennedy, Co Wicklow, Ireland (corresponding author, ted.lynch@coilte.ie).

associated yield curves and vice-versa by the mathematical operations of differentiation and integration.

A plethora of methods have been developed to represent forest growth in even-aged stands. These can be classified using a relatively small number of criteria that convey the essential differences. *Static* models assume that stands are managed to some prescribed pattern over the rotation. Within a European context the FC Management tables (Forestry Commission, 1981) is the best-known example. *Dynamic* models, on the other hand, do not assume prescribed management regimes and therefore can be used to forecast the outcome of a wider range of thinning practices. These models rely on modelling incremental changes in the variables of interest over time. The corresponding yield is then obtained by integrating or summing the incremental changes.

García (1979) gave a fillip to dynamic models through the development of systems of differential equations to model forest yield. The multivariate extension of the Bertalanffy-Richards model developed by García permit simultaneous estimation of growth component equations can take place. García introduced further ideas from systems theory into a forestry context, to describe how a forest stand evolves over time. This requires:

1. an adequate representation of the system (stand variables) at any point in time - the so called *state* of the system, and
2. estimates of the rate of change of state, and of the current value of any external control variables.

Using such a framework the central elements of differing yield prediction systems can be compared. To illustrate, models based on individual tree position maintain a more detailed state description compared with those that operate at the stand-level. From a practical forest management viewpoint, a growth model must provide the requisite forecasts without too much in the way of operating overheads. This requirement has led to growth models that are used for management purposes being developed at the aggregation level of a forest stand. Forest level prediction is then made through aggregating stand level predictions.

This work sets out by describing a class of functions that jointly constitute a García yield model. This class is augmented with a function for generating assortments. The fitting of these functions in the presence of sampling biases is subsequently described. General aspects relating to the deployment of the resulting dynamic yield model are discussed.

Methods

The development of stand-level growth models for Sitka spruce proceeded by adopting the state-space modelling methodology advocated by García (1984). That is, a small number of stand variables were chosen to represent the stand. Future states of the stand can then be determined from the current state, provided any future actions such as thinnings are detailed. The variables included in the state allow the subsequent calculation of quantities of interest such as stand volumes, log assortments and thinning reductions. Thus, a stand volume equation is employed to estimate stand volumes; an assortment generator predicts proportions of stand volume found in various log categories; and a thinning equation calculates basal area or stocking reductions during any possible thinning. A formal description of each component sub-model is now given.

Growth equations

The growth projection mechanism employed consisted of a system of differential equations formulated by García (1979). The system is sufficiently flexible to permit both empirical growth modelling, where no assumptions are made as to relations between stand variables, and modelling in limited data situations where known biological principles are used as a basis for forming equations (García and Ruiz 2003).

In formulating this system, García proceeded by way of generalising equation (1), the basic univariate Bertalanaffy-Richards model expressed as a linear differential equation in x^C. Here variable x is some measure of size such as top height and parameters are a, b and c.

$$\frac{dx^C}{dt} = ax^C + b \tag{1}$$

The power transform (2) is required to generalise the scalar exponential term x^C present in equation (1) to multivariate situations. The left hand side of (2) employs a non-standard mathematical notation, a vector being raised to a matrix power. This is computed using the expressions on the right hand side of the identity (assuming the logarithm and exponential functions are extended to a vector argument on an element-by-element basis).

$$\mathbf{x}^{\mathbf{C}} \equiv \exp(\mathbf{C} \ln \mathbf{x}) \tag{2}$$

Here $\mathbf{x}$ is an arbitrary n dimensional vector and $\mathbf{C}$ is an arbitrary $n \times n$ matrix. To illustrate, if the state vector used to represent the stand is chosen to be basal area (B) ($m^2\ ha^{-1}$), stocking (N) (stems ha^{-1}), and top height (H) (m), then take $\mathbf{x} = (B, N, H)^T$ and $\mathbf{C}$ as a 3 x 3 matrix of *reals* with elements c_{ij}, then by extending the exponential and logarithm functions as indicated above the resultant 3 x 1 vector can be expressed as

$$\mathbf{x}^{\mathbf{C}} = \begin{bmatrix} B^{c_{11}} N^{c_{12}} H^{c_{13}} \\ B^{c_{21}} N^{c_{22}} H^{c_{23}} \\ B^{c_{31}} N^{c_{32}} H^{c_{33}} \end{bmatrix}$$

The above illustrates that raising the state vector to a matrix power is a form of coupling transformation, since a new vector has been formed whose elements are multiplicative combinations of the original vector elements raised to appropriate powers.

The analogous multivariate system is specified as

$$\frac{d\mathbf{x}^{\mathbf{C}}}{dt} = \mathbf{A}\mathbf{x}^{\mathbf{C}} + \mathbf{b} \tag{3}$$

The above approach proceeds from the univariate to the multivariate model by way of exploiting the power transform given in equation (2). Within the differential equation system (3), a single equation is devoted to predicting top height growth. This equation also governs how the effects of site are incorporated into all of the system equations in (3). The top height development equation has equation form (1) which, assuming $H(t)$ denotes top height in metres at age t, may be expressed as

$$\frac{d\,H^C}{d\,t} = b(a^C - H^C) \tag{4}$$

Whose solution is

$$H(t) = a\left(1 - \left(1 - (H_0 / a\ )^c\right)\exp(-b(t - t_0))\right)^{1/c}$$

When parameter b is used as the site indexing parameter then the site index curves that arise are termed *polymorphic* and curves differ by what amounts to a *time scaling*. Parameter a is then interpretable as the asymptotic top height while parameter c can be regarded as a linearisation parameter, since equation (4) is a linear differential equation in the variable H^C. Equation (4) has been integrated subject to the initial condition $H(t_0)=H_0$.

From the solution of equation (4), b may be expressed in terms of the site index S (top height at age 30) and the stand top height initial condition

$$b = -\frac{1}{30-t_0}\ln\left[\frac{a^c - S^c}{a^c - H_0^c}\right] \tag{5}$$

Should a or c be chosen as the site indexing parameter then *anamorphic* site index curves would arise. Under a polymorphic site index curve the effect of site within the system (3) can be accommodated through scaled time. This means **A** and **b** change by a multiplicative factor to express site differences, while **C** is independent of site. Consequently, in system (3) there is no relationship between site and the state variables (García 1984).

Assuming the state vector has elements $\mathbf{x} = (B, N, H)^T$ then the following form of system (3) was employed to represent growth on an empirical basis for thinned and un-thinned stands. The derivatives are now taken with respect to $\tau = bt$, the scaled time variable.

$$\frac{dB^{C_{11}}N^{C_{12}}H^{C_{13}}}{d\tau}=a_{11}B^{C_{11}}N^{C_{12}}H^{C_{13}}+a_{12}B^{C_{21}}N^{C_{22}}H^{C_{23}}+a_{13}H^{C_{33}}+b_1$$

$$\frac{dB^{C_{21}}N^{C_{22}}H^{C_{23}}}{d\tau}=a_{21}B^{C_{11}}N^{C_{12}}H^{C_{13}}+a_{22}B^{C_{21}}N^{C_{22}}H^{C_{23}}+a_{23}H^{C_{33}}+b_2 \qquad (6)$$

$$\frac{dH^{C_{33}}}{d\tau}=-H^{C_{33}}+b_3$$

In terms of matrices **A** and **C** now have dimension 3 x 3, **b** is a 3 x 1 vector. García (1984) deployed a useful version of this model that arises from setting $a_{22} = a_{23} = b_2 = c_{21} = c_{23} = 0$. Then fitted values of c_{22} and a_{21} having opposite signs mean that stocking levels will decrease over time. Setting a number of parameters to zero also diminishes the number of parameters to be found. This simplified model was used as a base model prior to the fitting of system (6).

García (1984, 1989) also considered extensions to system (6) so as to examine responses to fertiliser treatments and thinning effects. Both can be examined using multiplier functions which are effectively time scaling devices. Thinning effects can also be modelled through augmenting the state vector through the use of an additional state variable R representing *relative closure* or site occupancy after thinning. Relative closure has values between 0 and 1. Assuming full closure is represented by a value of 1, its value after a thinning is reduced to the basal area of the thinned stand as a proportion of the basal area prior to thinning. After thinning, closure is assumed to increase towards its asymptotic value of 1. Relative closure is an unobservable variable in that its value is not directly available at all times through measurement, however, it is still amenable to inclusion within a state vector since the means by which its value changes are known.

System (6) is linear in the transformed state vector $\mathbf{x}^C$ and consequently can be integrated analytically between times that do not involve thinning. Given the initial state $\mathbf{x}_1$ at time t_1, García (1984) indicates the expected solution at time t_2 as

$$\mathbf{x}(t_2) = \left\{\mathbf{a} + \mathbf{P}^{-1}e^{\Lambda b(t_2-t_1)}\mathbf{P}\left[\mathbf{x}(t_1)^{\mathbf{C}} - \mathbf{a}\right]\right\}^{\mathbf{C}^{-1}} \qquad (7)$$

Where Λ (diagonal matrix of eigenvalues) and **P** (matrix of left eigenvectors) form the eigenvalue-eigenvector decomposition of the matrix $\mathbf{A} = \mathbf{P}^{-1}\Lambda\mathbf{P}$, and $\mathbf{a} = \mathbf{A}^{-1}\mathbf{b}$ is the asymptote vector for system (3). Note the presence of the time scaling parameter b within the exponential term - this represents the effect of site adjusting time.

Volume equation

A common stand level volume equation that predicts the volume basal area ratio as a function of top height is

$$V / B \cong a+bH \qquad (8)$$

Where a and b are parameters to be determined. The ratio V/B is used in that it tends to induce homogeneity in the error variance.

Beekhuis (1966) observed that equation (8) cannot be valid both before and after a thinning. Thinning *from below* tends to remove trees of smaller height so the volume to basal area ratio must increase after a thinning event. García (1984) suggested a modification whereby further terms involving N or B are admitted as predictors. Site index, S, may also be used to account for site effects (García, pers. comm.). A stand-level volume equation is then generally determined as

$$\frac{V}{B} \cong \beta_0 + \sum_{i=1}^{n} \beta_i \, g_i \left(B, N, H, S\right)$$

The right hand side of equation (9) denotes that the volume to basal area ratio is expressed as a linear function of terms involving basal area, stocking, top height and site index. Each g_i function is a multiplicative expression of its arguments that are possibly raised to powers. Explicitly, equation (9) is expressed in terms of a predictor set such as H, $H/\sqrt{N}$, NH/B, $1/H$, H/N, B/H, S/B, S/B^2 and the model identified by stepwise linear regression techniques.

Thinning equation

The thinning equation proposed by García (1984) allows for the determination of post-thinning basal area (*resp.* stocking) when the top height, pre-thin basal area, pre-thin stocking and post-thin stocking (*resp.* basal area) are known. When expressed as a differential equation it has the form

$$\frac{d \ln B}{d \ln N} = a \, B^b \, N^c \, H^d \qquad (10)$$

The left hand side of (10) is interpretable as the percentage change in basal area arising from a percentage change in stocking. This interpretation follows from the expression $d \ln B / d \ln N = (dB/B)/(dN/N)$. Consequently, equation (10) is a model of the stocking elasticity of basal area (Silberberg 1990).

The equation is separable and on integration its solution is obtained as

$$\ln B = -\frac{1}{b} \ln \left[B_0^{-b} - \frac{ab}{c} H^d \left(N^c - N_0^c \right) \right]$$

Where H denotes the top height, B_0 and N_0 are the pre-thin basal area and stocking and the post-thin values are B and N. The model parameters to be determined by nonlinear least squares are a, b, c, and d.

A useful property of differential equation models is that they are closed under the operation of *composition*. With respect to the thinning equation this means a thinning removing 200 stems ha^{-1} followed immediately by another thinning removing 300 stems ha$_{-1}$ results in the same basal area reduction as a thinning of 500 stems ha$_{-1}$. This consistency property will always arise when differential equations are used as transition functions. System (3) gives rise to a similar consistency condition when projecting over time (García 1994).

Using the thinning equation (11) in conjunction with the stand volume equation (9) allows for thinnings to be performed by volume reduction. The thinning equation permits the stand volume after thinning to be expressed as $V(B, N(B; B_0, N_0), H)$, where post-thin stocking is given as a function of post-thin basal area (pre-thin basal area and stocking appear as parameters within the thinning equation). This means Newton's algorithm (Conte and de Boor 1972), or some similar technique for locating function roots, can be deployed to reduce the basal area from its pre-thin level, to post-thin level, to simulate the extraction of a specified volume of thinnings. Similarly, it is possible to simulate the effect of thinnings by increasing quadratic mean diameter. The procedure would be to employ a root finding technique to reduce B and at each step, use the thinning equation to predict $N(B; B_0, N_0)$ and consequently recover the quadratic mean diameter (terminate the step-wise procedure with success if the quadratic mean diameter has reached its target value).

Assortment equation

An assortment model is used to calculate assortments for any production thinnings and the crop. Generally, the problem of estimating assortments is that of delineating a set of categories (log classes), and determining the proportion associated with each category (assortment), such that the proportions sum to unity. Past endeavours to solve this problem have focused on sectioning a diameter distribution (García 1981) and subsequently employing a taper function to calculate the quantities of interest. More recently, the problem has been addressed using techniques that model the proportions directly, such as multinomial response models (Arabatzis and Gregoire 1990). These are a class of what are known as limited dependent variable regression models.

The model defined as Unordered MultiNomial Logistic (UNML) by Arabatzis and Gregoire (1990) was employed here largely because of its flexibility in permitting differing numbers of predictors to be associated with each *category*, and because it does not require any common slopes assumption as does the Ordered MultiNomial Logistic (OMNL) model. Assuming a variable Y is used to indicate the J+1 categories, indexed 0 through J, then the UNML model is expressed as

$$\begin{aligned} \Pr\left[Y_i = j\right] &= \frac{\exp\left(\hat{\mathbf{a}}'_j \mathbf{x}_{ij}\right)}{1 + \sum_{k=1}^{J} \exp\left(\hat{\mathbf{a}}'_k \mathbf{x}_{ik}\right)} \\ \Pr\left[Y_i = 0\right] &= \frac{1}{1 + \sum_{k=1}^{J} \exp\left(\hat{\mathbf{a}}'_k \mathbf{x}_{ik}\right)} \end{aligned} \qquad (12)$$

Here i denotes an observation subscript and j a category subscript. While $\mathbf{x}_{ik}$ denotes predictors associated with category k and observation i. Finally β_k denotes parameters for category k. The probability of category 0 is determined by subtraction.

The equations given in (12) can be adapted to obtain assortment predictions for crops with, or without, thinning. The approach involves partitioning the predictors associated with each category (the observation subscript has been dropped for convenience).

$$\mathbf{x}_j^T = \mathbf{x}_{j1}^T \mid \mathbf{x}_{j2}^T \qquad (13)$$

Where

$\mathbf{x}_{j1}$ is the set of predictors for category *j* associated with the main crop; and
$\mathbf{x}_{j2}$ is the set of predictors for category *j* associated with the thinning.

The underlying idea is that an assortment for a thinning is obtained by way of a modification to the pre-thin crop assortment. The extent of the modification to the crop assortment depends on the severity of the thinning. The predictor variables in the vector $\mathbf{x}_{j2}$ form part of the prediction only if a thinning assortment is being predicted (they are defined to be zero otherwise). Generally, variables in $\mathbf{x}_{j2}$ have special structure that provides a measure as to the extent of thinning. This is achieved by ensuring they are functions of $N_0 - N_1$, $B_0 - B_1$, $D_1 - D_0$. Here N_0 and N_1 are the stocking per hectare before and after thinning, B_0 and B_1 the respective basal areas, and D_0 and D_1 respective quadratic mean diameters (expressed in centimetres). Ratio terms can also be utilised.

Predictions are made for two log categories. Category one is an estimate of the proportionate volume of logs of at least 3 m length, and having small end diameter (SED) equal to or greater than 20 cm. Category two estimates the proportionate volume of logs having small end diameter in the interval 14 to 20 cm. Category 3, is obtained by difference from unity, and estimates the proportionate volume in the 7-14 cm SED category.

The standard multinomial model has been modified in two respects. Firstly, categories are permitted to have a different numbers of predictors associated with them. Secondly, a weighted form of the likelihood used in parameter estimation has been formulated to cope with data imbalances (see Appendix 1). Using a different number of predictors for each category means good model identification can take place via a log odds ratio and stepwise linear regression techniques.

Data provenance

The bulk of the data used in fitting models were extracted from Coillte Teoranta's (the Irish Forestry Board – the state commercial forestry company) permanent sample plot record system. The associated database contains records from many silvicultural and thinning trials established during the period 1963 to 2001. The trials were initially established as replicated experimental designs with repeated measurement. Issues relating to the use of *blocked data* within a yield modelling context are discussed by Broad and Lynch (2006).

Aggregating data across repeatedly measured designed experiments, each of which is a well-designed experiment, does not guarantee a well-structured data set for the purposes of yield modelling. Coillte has no experiments established in un-thinned stands having site indices below 10.5 m, yet some 21.1% of its un-thinned estate has site indices below this level. Similarly, for thinned stands with site indices below 13.5 m, approximately 1.2% of its experimental data falls in this range compared with 21.16% for the estate data (see Figure 1, Broad and Lynch (2006)).

These data omissions would be best rectified through an additional sampling scheme specifically designed to address the imbalance. The approach taken here was to use data substitution to facilitate growth model development in the absence of further sampling. Substituted data occurred below site index 15.0 (m) were sourced from Booklet 48, Forestry Commission (1981).

Parameter estimation

The above equations were fitted to data extracted from the Coillte permanent sample plot database. Obtaining data records in a form suitable for model fitting was facilitated by a database query program written using Microsoft Access®. The program forms a useful tool for validating data from permanent sample plot records and subsequently extracting data for growth modelling purposes. The validation code performs both basal area, and sample tree volume checks. The data extraction code takes the validated records and prepares aggregate information for growth modelling purposes.

Weighting techniques were used to address some of the data imbalances within Coillte's research database (Broad and Lynch 2006). The data imbalances include unrepresentative sampling in some site index classes. The procedure for determining weights to be used in estimation procedures relies on knowledge of some attribute variable for both the sampled population and the target population. The attribute variable chosen in this study was site index. Once site index histograms are available for the sampled and target populations, sampled observations can be weighted as follows.

$$P_T \equiv \left(\frac{P_T}{P_S}\right) P_S \qquad (14)$$

Where P_s denotes the proportion of sampled observations in a specified site index class and P_T is the portion of target observations in the same site index class. The sampled observations are weighted by the quotient ratio in parentheses to ensure a weighting consistent with the target population. This weighting scheme is applicable only when observations exist in the sampled site index class (they must then exist in the target class). If the target class has observations but the sampled class does not then it is necessary to aggregate over classes to ensure sampled observations exist within the aggregated sample class.

When determining the weights for a specific equation the histograms for the target and sampled populations are initially constructed. For the target population the histograms always used in weight determination are those indicated as *Estate Un-thinned* and *Estate Thinned* in Figure 1 of Broad and Lynch (2006). For the sampled population the histograms are constructed from the data used in equation fitting and consequently vary between equations. The sampled population histograms are constructed after any additions for data omissions have taken place. This process may not uniquely define a set of weights, as often more than one possibility exists for drawing the sampled population histogram. To illustrate, when fitting growth component models the sampled population histograms are constructed from the trajectory data used in model fitting. Different histograms usually arise depending on whether we use the plot site index or trajectory site index as the observation when constructing the histogram. Only when all plots have the same number of trajectories are the histograms the same.

Height growth estimation

García (1983) assumed in estimating the parameters of the height equation (4) that its increments were perturbed by a *stationary Gaussian process* with independent increments (see also Seber and Wild 1989). This permits construction of the log likelihood function and determination of the height equation parameters through log likelihood maximisation.

Site index curves are typically obtained by varying one parameter from plot to plot, although strictly speaking any function of the parameters can be chosen to vary in this manner (García 1983). Parameter(s) chosen to represent site are referred to as local parameters within the likelihood function, while those that are constant across all sites are termed global. Both anamorphic (a local) and polymorphic (b local) height equations were fitted to a height data set extracted from Coillte's database augmented with FC height trajectory data to cover data omission in lower site index classes (Broad and Lynch 2006). The final data comprised 423 trajectories of which 160 were from FC data. Weighting changes for the log likelihood were achieved by replicating observations. The FC data was afforded a higher weighting over repeated fitting attempts to lower predictions for lower site index classes.

The log likelihood function for the polymorphic model had a value of 3887.412. This was an improvement over the anamorphic model which was not considered further. The use of the polymorphic model means the suggested time scaling in system (6) can be implemented.

In the solution of the height equation the initial height H_0 (m) at time t_0 (years) is assumed to be zero. The following parameter estimates were obtained using García's height estimation program

$a = 49.348040$
$c = 0.624157$
$t_0 = -0.75$

The t_0 parameter was pinned although various candidate values were considered.

The height equation can be used to predict top heights for trees in production environments. There are non-production areas in Ireland where this equation would be conservative. Figure 1 shows a family of site index curves.

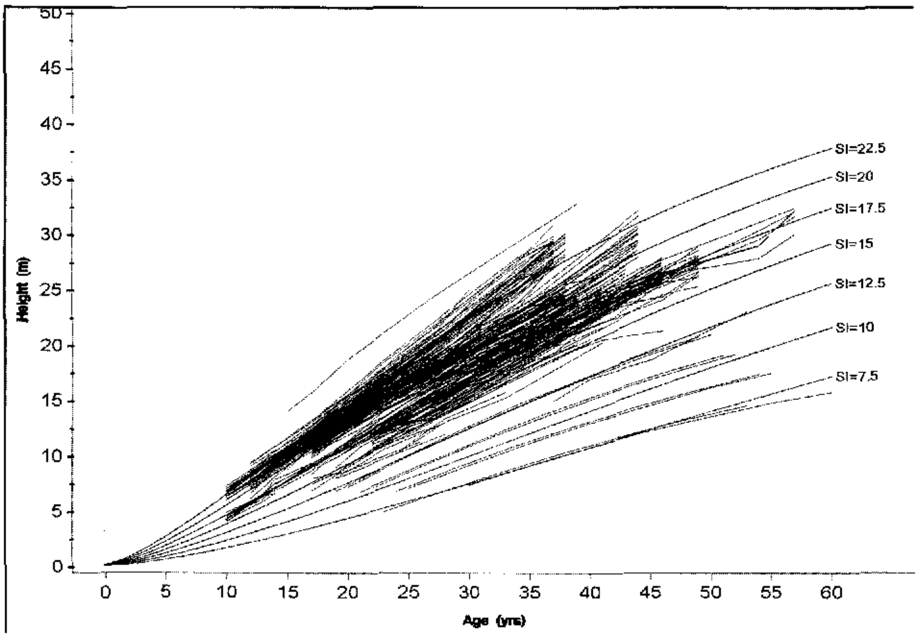

Figure 1: Sitka spruce site index curves (reference age 30 yrs).

Basal area and mortality growth estimation

Consecutive stand measurements of basal area, stocking and top height (state vector elements) in the absence of thinning form the basis of the data set used to fit the dynamical system specified in equations (3). In a thinned stand, a *growth trajectory* comprises the stand state measurement immediately after thinning (post-thin) along with that immediately prior to the subsequent thinning (pre-thin). In un-thinned stands a trajectory will arise from each consecutive pair of stand measurements. In thinned stands, windthrown trees that occur immediately prior to thinning were modelled as part of the thinning. This allows for limited recoverability of windthrown material.

The thinned model has been fitted to trajectories from stands having had *selection*, *systematic*, or *line plus selection* thinning treatments. A limited number of the trajectories from wider spaced stands were used. In addition, trajectories from un-thinned stands were incorporated in the thinned data set. This reflects the fact that all stands start out as un-thinned stands and that a wide diversity of thinning practice can exist. Additional FC data were added to address the issue of data being absent from lower site index classes (see Broad and Lynch 2006) Careful scrutiny was given to screening plot data. Basal area and mortality increments were examined along with a graphical analysis of trajectory data. In cases of obvious abnormalities, these data were no included.

For parameter estimation purposes, García (1984) assumed that the system of differential equations (3) is perturbed by a 3-dimensional Wiener-process. The resulting system of stochastic differential equations facilitates the construction of the conditional probability of observing the stand state at time t_2 given the state at time t_1. The assumption of statistical independence between plots enables the construction of the log likelihood function that provides a probability model of generating the observed data. The log likelihood function may be estimated under either a diagonal correlation structure (Method I) or full correlation structure (Method II). Method I was used throughout, as previous applications indicate that it is the most suitable for estimation purposes. The model parameters were obtained by maximisation of this log likelihood function. The full derivation of the log likelihood function is given in García (1979) and a modification to allow for zero eigenvalues is stated in García (1984). The case of zero eigenvalues is consistent with an over-specified state vector for growth modelling. However, such a vector may be useful in that its components may be used as arguments within ancillary functions such as the volume or thinning equations.

In response to the data imbalance issues identified in the research data (Broad and Lynch 2006) the weighted form of the Method I log likelihood function was developed. This was coded using Compaq FORTRAN (Compaq Computer Corporation, 2001). In all fitted models, weights were initially determined using equation (14). For each fitted model the sampled population histogram was constructed using the site indices associated with the trajectory data. Thereafter weights were modified to ensure *ad hoc* criteria were satisfied. These criteria were designed to ensure that relativities between thinned and un-thinned stands being realistic, projections across the range of site indices remaining reasonable and comparisons with FC projections remained satisfactory.

Table 1 indicates the structure of some models fitted subsequent to fitting the top height equation. For thinned and un-thinned stands the base model with parameters a_{11}, a_{12}, a_{13}, a_{21}, c_{11}, c_{12}, c_{13}, c_{21}, b_1 was initially fitted - these are models 1(T) and 4(U) in Table 1. Subsequently, the fully parameterised versions of system (6) were fitted - these are models 2(T) and 5(U) in Table 1. When examining projections from these

***Table 1**: Model fitting results (model number & thinning status (T= thinned, U=un-thinned), and parameter list (excluding top height parameters)).*

Model	*Log likelihood*	*Number of trajectories*	*Model description*
1(T)	*7926.91*	*1460*	$a_{11}, a_{12}, a_{13}, a_{21}, c_{11}, c_{12}, c_{13}, c_{21}, b_1$
2(T)	*8017.02*	*1460*	$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, c_{11}, c_{12}, c_{13}, c_{21}, c_{22}, c_{23}, b_1, b_2$
3(T)	*8028.80*	*1460*	$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, c_{11}, c_{12}, c_{13}, c_{21}, c_{22}, c_{23}, b_1, b_2, \gamma_1, \gamma_2$
4(U)	*12560.6*	*1265*	$a_{11}, a_{12}, a_{13}, a_{21}, c_{11}, c_{12}, c_{13}, c_{21}, b_1$
5(U)	*12763.4*	*1265*	$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, c_{11}, c_{12}, c_{13}, c_{21}, c_{22}, c_{23}, b_1, b_2$
6(U)	*12888.8*	*1265*	$a_{11}, a_{12}, a_{13}, a_{21}, a_{22}, a_{23}, c_{11}, c_{12}, c_{13}, c_{21}, c_{22}, c_{23}, b_1, b_2, \gamma_1, \gamma_2$

models it was noted that for lower site indices the fitted models tended to over-estimate when comparisons were made against FC tables. This led to an investigation of the site scaling method in system (6).

The method deployed to investigate site scaling was to use the multiplier function approach of García (1989) adapted to the scaling of the polymorphic height parameter b, rather than *relative closure* as occurs in García's application. The method was based on using a function $g(b)$ to site-scale the first two equations in system (6) rather than b. The requirements of such a function are that it be defined on some domain where $g(b)$ is less than b, thereafter it has value b. Several candidate functions were tried, the most successful of which was

$$g(b) = \begin{cases} b\left(1 + \gamma_1 (b - \gamma_2)^2\right) & \text{if } b < \gamma_2 \\ b & \text{otherwise} \end{cases} \tag{15}$$

Equation (15) is a spline comprising a modified cubic followed by a linear portion. The parameter γ_2 denotes the join point at which $g(b)=b$ and $g'(\gamma_2)=1$, the additional condition is $g(0)=0$. Nonlinear site scaling is attained for values of b below γ_2. When fitted, the parameter values obtained allow a zero for $g(b)$ at a small positive value of b. These values correspond to site indices that are outside the range of the fitted model and should pose no threat to projections.

Augmenting the fully parameterised system (6) with equation (15) led to models 3(T) and 6(U) in Table 1. All the thinned models in Table 1 have the same observation weight set. Similarly, all un-thinned models have a common, but different, weight set. From the increases in the log likelihood function at optimality in Table 1, it is apparent that the nonlinear site scaling has a beneficial effect for both thinned and un-thinned stands. Parameters for the fully parameterised models with augmented site scaling can be found in Box 1 & Box 2 of Appendix 2.

Explanation of the need of nonlinear site scaling could include: genuine nonlinear site effects at lower site indices with respect to basal area and stocking development; the data interface between Forestry Commission and Irish research data requiring nonlinear site scaling to rectify site response imbalances; an ill-fitting height equation. Discriminating between these possibilities is difficult but it is suspected that non-linear site scaling to rectify site response imbalances is responsible. The value for γ_2 corresponds to site index 14.35 (m) for thinned stands and 17.34 (m) for un-thinned stands.

Plots of observed (also used in model fitting) *v* predicted trajectories are illustrated in Figures 2-6.

Figure 2 also serves to illustrate the range of data used in the fitting of the thinned model, and gives an indication of the regions of the state space within which model projections can be made. The dense cloud of observations less than age 45 represents the bulk of the research data.

Figure 3 illustrates stocking development in thinned stands, with the bulk of the data with stocking levels of 3500 stems ha^{-1} and less.

Height development is indicated in Figure 4. Both thinned and un-thinned models share a common height equation.

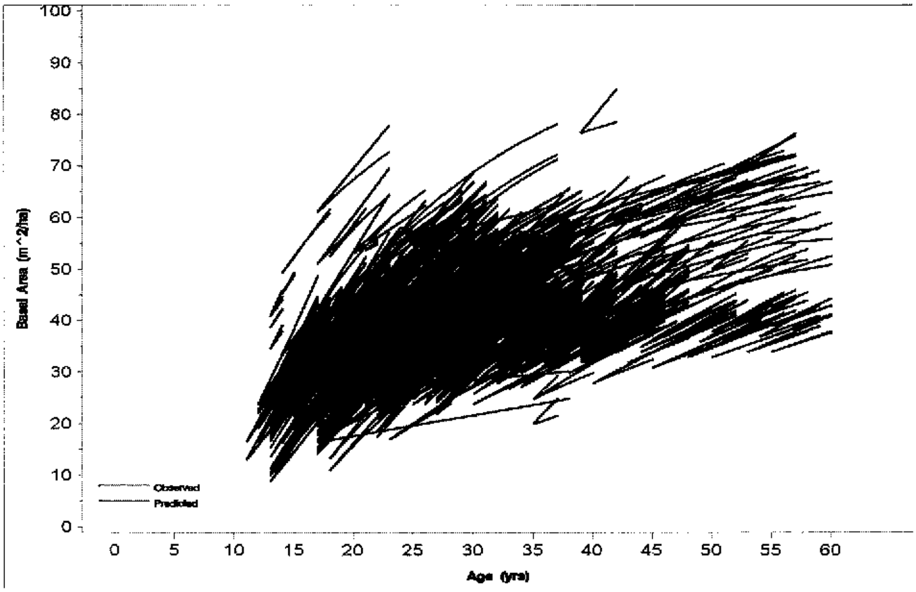

Figure 2: Sitka spruce (thinned) observed v predicted basal area development.

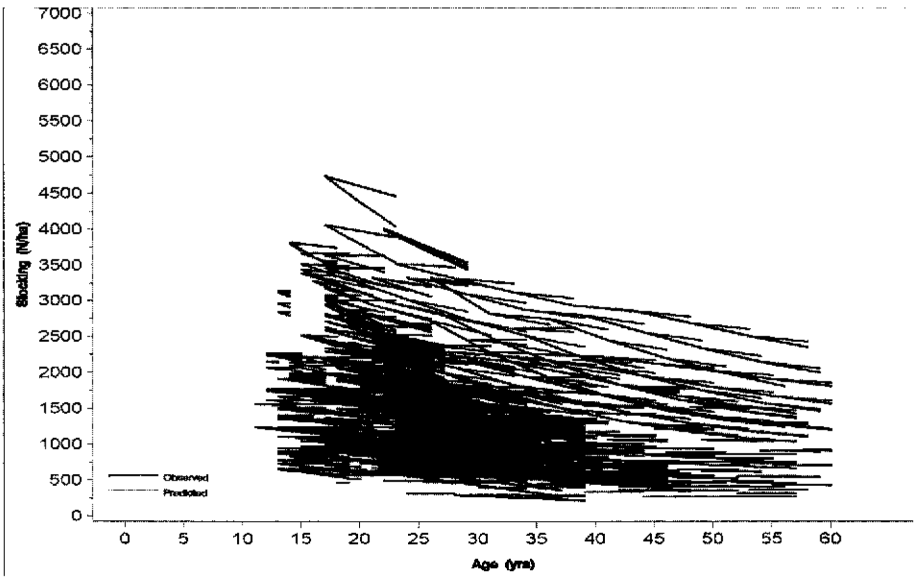

Figure 3: Sitka spruce (thinned) observed v predicted stocking development.

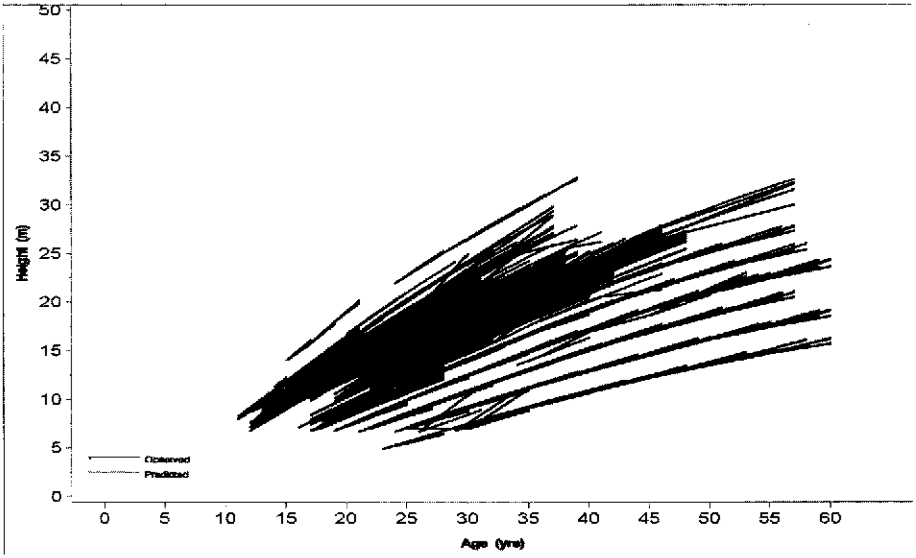

Figure 4: Sitka spruce (thinned) observed v predicted top height development.

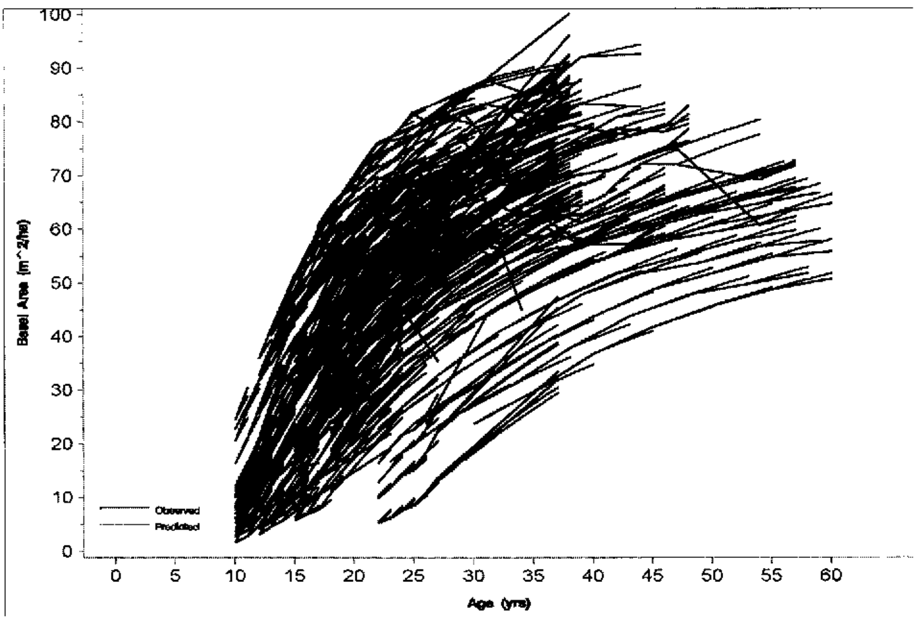

Figure 5: Sitka spruce (un-thinned) observed v predicted basal area development.

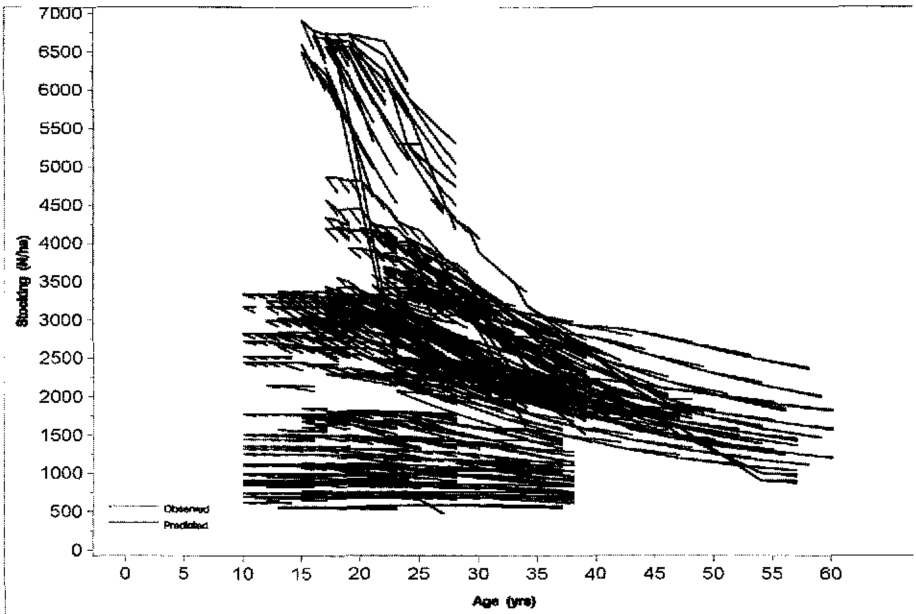

Figure 6: Sitka spruce (un-thinned) observed v predicted stocking development.

From Figures 5 & 6 it is immediately apparent that the observed basal area and stocking development from thinned stands is considerably less variable than that from un-thinned stands. Mortality in un-thinned stands tends to be clumped both spatially, and temporally.

Volume equation estimation

The explicit form of equation (9) identified by SAS® Proc Reg (SAS Institute 1990) using stepwise regression was found to be

$$V/B = \beta_0 + \beta_1 H + \beta_2 H/\sqrt{N} + \beta_3/H + \beta_4 H/N + \beta_5 100/(BH) + \beta_6 S/B + \beta_7 S^2/B$$

Where
$\beta_0 =$ 3.30847 (0.1415)
$\beta_1 =$ 0.32046 (0.0106)
$\beta_2 =$ -0.53576 (0.2815)
$\beta^3 =$ -21.31789 (0.0001)
$\beta_4 =$ 23.29623 (0.0244)
$\beta_5 =$ 0.77977 (0.1671)
$\beta_6 =$ -3.05454 (0.2538)
$\beta_7 =$ 0.12462 (0.0086)
with adjusted $R^2 = 0.9806$.

Multiplying the above equation by basal area, B, puts it in a suitable form to predict volumes for both thinned and un-thinned stands.

The equation was fitted to a dataset comprising 1676 observations that was extracted from Coillte's database. Each observation had measurements of basal area, stocking, top height, site index and volume to 7 cm small end diameter (*SED*). Some 113 observations came from the non-research database. This was used as a calibration dataset, as it is known to be free of sampling bias. A further 222 observations comprised replicated FC data to cover the shortfall of data in lower site index classes. The remaining observations came from the research database and a small external dataset known as the *paper plot* series.

The fitting strategy commenced with weight determination using equation (14). The site index available with each observation was used to determine the sampled site index distribution. Upon fitting, the relative errors with respect to the calibration data were determined. Subsequent refitting involved iterative weight adjustment in order to reduce the relative prediction errors with respect to the calibration data. The fitting procedure involved a manual search to determine weight sets that were associated with low relative prediction errors over the calibration data. The approach of treating the weights as extra parameters and using nonlinear programming techniques for parameter estimation could not be used. This approach would have tended to drive the weights associated with the research data to zero. The research data were observed over a wider range than the calibration data, and can be seen to extend the range of the calibration data, despite the quality issues.

Thinning equation estimation

The parameters in equation (11) were determined using the SAS® weighted non-linear least squares procedure, Proc Nlin (SAS Institute, 1990). The data comprised 532 observations each having basal area and stocking (before and after thinning) along with the top height. Observations from plots with systematic first thinning were excluded from the regression. Observations were assumed to be statistically independent. The model represents the *pure selection component* of any thinning (i.e. exclusive of any rack-row thinning). The fitting process was direct, as Coillte does not have any thinning data independent of those used in model fitting. Consequently, there was no basis for undertaking further model fits through weight changes.

The weights for the nonlinear regression were initially determined using equation (14). Site indices for the sampled population were obtained by union of the site indices from each thinning observation. Data was sourced from the portion of Coillte's database representing stands over a range of thinning types: light selection, systematic, selection, spaced thinnings. Additional data came from an early thinning data set representing non-systematic thinnings.

The estimated thinning equation parameters and their parenthesised standard errors are

$a = 3.140$ (0.7122)
$b = 0.1800$ (0.0414)
$c = -0.1838$ (0.0253)
$d = -0.3267$ (0.0463)

During simulations, rack-row thinning is dealt with using proportionate reduction of basal area and stocking.

Assortment equation estimation

Equation (12) was used to develop separate assortment models for thinned and un-thinned Sitka spruce stands. Parameter estimation was through maximum likelihood estimation. The likelihood function specified by Greene (1997, p. 916) was extended to the weighted form specified in Appendix 1, so as to address imbalances in volume data collection (Broad and Lynch 2006). The likelihood was maximised using the Newton-Raphson method, which functions through repeated solution of a set of linear equations. A useful property of the log likelihood function is that it is globally concave, thereby ensuring that a local optimum is a global one, and greatly facilitating the estimation of its parameters (Greene 1997). Code development was via the SAS® Interactive Matrix Language (SAS Institute 1989).

Fitting of the log-odds ratio via weighted least squares was undertaken prior to maximum likelihood estimation. This offers the advantage of furnishing starting estimates for maximum likelihood estimation. Furthermore, the diagnostics available within regression packages can be used for model identification purposes.

The assortment model for thinned stands has some 384 thinned assortment observations that were aggregated over systematic- and line-selection thinnings from research plots. An additional 70 observations came from the non-research data and the remaining 112 observations were FC data to cover the data omission with respect to lower site indices. Final parameters for the fitted model are given in Box 3 of Appendix 2. The assortment model for un-thinned stands has 522 assortment observations from research plots, 43 observations from non-research plots and 113 additional observations sourced from FC data (Forestry Commission 1981).

The strategy adopted in fitting was to calibrate with respect to the non-research data as this data set is considered bias free. The research volume data is considered to be biased due to measurement problems associated with establishing the volume/basal area regression (Broad and Lynch 2006). Consequently, weights that were initially established using equation (14) were modified on an *ad hoc* basis so as to reduce relative prediction errors with respect to the calibration data and the fitting repeated. The fitting procedure was terminated when a mean relative prediction error of 0.0065 was achieved for the calibration data. The weight modification strategy is essentially a search, with manual intervention at each iteration.

Discussion

The most serious difficulty encountered during the construction of these models was how to address the sampling biases present within the data. Broad and Lynch (2006) indicate that only one of these biases is specific to the experimental design data. Biases were introduced when the repeated measures data from replicated experiments were aggregated and the resulting panel data set considered for yield modelling purposes. The biases are a consequence of the different data requirements for yield modelling data – notably the randomisation at plot level.

The volume bias is a genuine measurement bias. In extreme cases it leads to ill-conditioning of the volume basal versus area regression line. In these cases the bias was detected at the plot level during routine validation. In these instances correction was undertaken by creating a wider volume sample diameter range through aggregating data across blocks. The full significance of the volume bias was not however appreciated until the bias of the stand level volume equation was further investigated (Broad and Lynch 2006). The use of independent calibration data, while still retaining the biased data because of its beneficial spread, and subsequently using weight adjustment during fitting has proved a suitable way to address the bias. Both volume and assortment equations were fitted using this approach.

The building of the growth component models required addressing data omission, sampling imbalances due to over and under-sampling in some site index classes, and the compromising of statistical independence between blocked plots. Only the first two of these were addressed through data addition from an external source and by employing weighted versions of García's estimation techniques. The eroding of the statistical independence between plots could be halted, for example, by using only one plot from each experiment – this is hardly a practical solution however, as it would involve not using most of the data. Another alternative solution would be to investigate error-component models to deal with the inter-plot cross-correlation. Such models are not well developed and results from using them on even well-structured data are inconclusive (Gregoire 1987).

The extant FC models were formed by aggregating stand statistics over rigidly managed stands. As such they are largely free of equation error and provided invaluable references during model construction. Use was made of them not in terms of the absolute values of their projections but rather in terms of the relativities they offer for thinned and un-thinned stands. In addition, the trend information available on differences between stands of varying site index acted as a valuable resource.

The use of top height at a specified reference age (site index) to assess site potential is based on the observation that top height development is little affected by changes in stand density through either thinning from below, or initial spacing levels. Moreover, top height development is largely independent of the timing of such silvicultural operations and consequently its adoption as a mechanism to classify site potential is widespread. By way of comparison, the related concept of yield class (m^3 ha^{-1} yr-) employed by Johnston and Bradley (1963) has strong temporal constraints in that all thinning operations are assumed as having been performed on time. Consequently, the advancement or delay of any thinning means that yield class does not conform to its usage within the FC tables.

Although the height equation is perfectly adequate for modelling top height development, the interaction of neutral (or line thinning) with top height development requires consideration. In neutral thinning, trees are removed in proportion to their relative abundance within the stand. Thus, trees that would otherwise form part of the top height determination are removed and consequently top height may be reduced immediately following a neutral thinning. This effect could also be accompanied by a subsequent loss of height growth due to the manner in which the stand is opened. Both effects could be investigated through modelling. However, under comparatively low neutral thinning intensities the influence of both effects on top height development is likely to be small. Consequently, no modifications to the top height equation have been considered.

The classification of site effects using either site index or yield class are not totally unrelated. The use of yield class to assess site quality is appealing as it is also a measure of volume productivity. Because it is difficult to measure cumulative volume production at a given age, yield class is more easily obtained through measuring a strongly correlated variable - top height. Therefore, the same primary index - top height growth - is used in both systems. The FC models are based on the observation by Eichhorn in 1904 (Assmann 1970) that the stand volume to stand height relationship is independent of site. This means differential equations could be formulated for growth projection using top height as the independent variable rather than age as was done in this work. This would overcome the static element of the FC approach.

The flexibility of dynamic models in representing stand management scenarios comes at the price of maintaining sufficient mensurational data to form the starting point (state vector) for growth projection. The FC alternative of ensuring that establishment and subsequent thinning take place to a prescribed pattern places considerable restrictions on how stands can be managed. There may also be a large opportunity cost associated with the deployment of static yield models as strict adherence to them constrains the decision set available to managers. The application of a forest planning tool should not restrict the range of silvicultural options available to managers, particularly when those decisions can have strong economic consequences. In the case of a growth model, projections should not be predicated on a particular form of stand management.

Forestry Commission models give great weight to volume maximisation as a recommended or even optimal form, of stand management. Testifying to this is the defining of yield class and marginal thinning intensity, in terms of maximum mean annual (volume) increment (MMAI). For stands producing a single valuable resource, rotation length under volume maximisation is generally longer than that under optimal economic management (assuming repeated rotations). Equality holds only if the discount rate is zero. The management of repeated rotations under optimal economic management is equivalent to maximising the present value of an annuity payable at the end of each rotation (Neher 1990). Consequently, financial returns to forest growers can usually be improved by opting for some form of optimal economic management rather than volume maximisation. Further, dynamic growth models are readily utilisable within stand- or forest-level optimisers that seek optimal economic management decisions under changing cost and revenue structures. This in turn introduces substantial freedom as to how stand- and forest-level management activities can be conducted.

Acknowledgements

Thanks are due to Professor Oscar García (University of Northern British Columbia) who greatly facilitated the development of these models by his unstinting response to queries. Thanks also to Dermot O'Brien for his courage in entrusting part of the job to a lad from the far-flung isles. We acknowledge Myles MacDonncadha and Ciaran Doyle for their programming work. The authors acknowledge a referee. Funding for the work reported here was provided by COFORD and Coillte.

References

Arabatzis, A.A. and Gregoire, T.G. 1990. Ordered and unordered multinomial response models: an application to assess loblolly pine merchantability. *Can. J. For. Res.* 21:261-268.

Assmann, E. 1970. *The Principles of Forest Yield Study* (translated from the original German, Waldertragskunde, by Gardiner, S.H.). Pergamon Press, Oxford.

Beekhuis, J. 1966. *Prediction of yield and increment in* Pinus radiata *stands in New Zealand.* FRI Technical Paper No. 49. New Zealand Forest Service.

Forestry Commission. 1981. Booklet 48, *Yield Models for Forest Management.* Forestry Commission, Edinburgh.

Broad, L.R. and Lynch, T. 2006. Panel data validation using cross-sectional methods. *Ir. For.* (63).

Clutter, J.L. 1963. Compatible growth and yield models for loblolly pine. *For. Sci.* (9):354-371.

Compaq Computer Corporation. 2001. Compaq Visual FORTRAN, V6.6.

Conte, S.D. and de Boor, C. 1972. *Elementary Numerical Analysis – an algorithmic approach,* 2nd ed. McGraw-Hill, Kogakusha.

García, O. 1979. Modelling stand development with stochastic differential equations. In Elliott D.A. (Comp.) *Mensuration for management planning of exotic forest plantations*, p 315-333. FRI Symposium No. 20. New Zealand Forest Service.

García, O. 1981. Simplified method-of-moments estimation for the Weibull Distribution. *N.Z. J. For. Sci.* 11:304-306.

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 Down's forest. *N.Z. J. For. Sci.* 14:65-88.

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

García, O. 1994. The state-space approach in growth modelling. *Can. J. For. Res.* 24: 1894-1903.

García, O. and Ruiz, F. 2003. A growth model for eucalypt in Galicia, Spain. For. Ecol. & *Mgmt.* 173:49-62.

Greene, W.H. 1997. *Econometric Analysis*, 3rd ed. Prentice-Hall, New Jersey.

Gregoire, T.G. 1987. Generalized error structure for forestry yield models. *For. Sci.* 33:423-444.

Johnston, D.R. and Bradley, R.T. 1963. Forest management tables. *Comm. For. Rev.* 42:217-227.

Joyce, P. and OCarroll, N. 2002. *Sitka spruce in Ireland.* COFORD, Dublin.

Neher, P.A. 1990. *Natural Resource Economics Conservation and Exploitation.* Cambridge Univ. Press, Cambridge.
SAS Institute. 1989. *SAS/IML® Software,* 1st ed., v6.0.
SAS Institute, 1990. *SAS/STAT® User's Guide,* Vol. 2, 4th ed., v6.0.
Seber, G.A.F. and C.J. Wild. 1989. *Nonlinear Regression.* Series in probability and statistics. Wiley, New York.
Silberberg, E. 1990. *The Structure of Economics: A Mathematical Analysis,* 2nd ed. McGraw-Hill, Boston.

Appendix 1

Growth model

Trajectory data consists of consecutive stand measurement pairs of the form $(\mathbf{x}_1,\mathbf{x}_2)$. The scaled time interval between observations constituting a measurement pair is denoted by $\delta\tau$.

The weighted likelihood function has form

$$L = \prod_{j=1}^{n} \left[f\left(\mathbf{x}_{2j} \mid \mathbf{x}_{1j}\right)\right]^{w_j} \tag{A1}$$

where

$f(\mathbf{x}_{2j} \mid \mathbf{x}_{1j})$ is the conditional density of $\mathbf{x}_{2j}$ given $\mathbf{x}_{1j}$

w_j is the weight associated with the jth measurement pair.

Taking logarithms and substituting from Appendix expression (A2) (García 1984) and reworking García's Method I estimation technique eventually leads to the weighted log-likelihood function

$$\ln L = -\frac{1}{2}\left[\sum_{j=1}^{n} w_j\left(p\ln 2\pi + p + \sum_{i=1}^{p}\ln\hat{\sigma}_{ii}^2\right) + \sum_{j=1}^{n}\sum_{i=1}^{p} w_j \ln R_i(\Delta\tau_j)\right] \tag{A2}$$

$$+ \sum_{j=1}^{n} w_j \ln abs\left(|\mathbf{P}\|\mathbf{C}|\right) + \mathbf{1}'\sum_{j=1}^{n} w_j \ln \mathbf{x}_{2j}^{C} - \mathbf{1}'\sum_{j=1}^{n} w_j \ln \mathbf{x}_{2j}$$

where

$$\hat{\sigma}_{ii}^2 = \frac{\sum_{j=1}^{n} w_j \mathring{\mathbf{a}}_{ij}^2 / R_i\left(\Delta\tau_j\right)}{\sum_{j=1}^{n} w_j} \tag{A3}$$

All other notation is as specified in García's (1984) paper. Setting the weights to unity recovers García's Method I likelihood.

Assortment model

A multinomial Logit model was used as the basis for modelling assortments. The standard model has been extended to allow for weighting of observations and to permit a different number of predictors to be associated with each category.

The weighted likelihood function has form

$$L = \prod_{i=1}^{n} \left(\prod_{j=0}^{J} \Pr[Y_i = j]^{d_{ij}} \right)^{w_i}$$

where

$$\Pr[Y_i = j] = \frac{\exp(\hat{\mathbf{a}}'_j \mathbf{x}_{ij})}{1 + \sum_{k=1}^{J} \exp(\hat{\mathbf{a}}'_k \mathbf{x}_{ik})} \qquad j = 1, \cdots, J$$

$$\Pr[Y_i = 0] = \frac{1}{1 + \sum_{k=1}^{J} \exp(\hat{\mathbf{a}}'_k \mathbf{x}_{ik})}$$

$$\sum_{j=0}^{J} d_{ij} = 1$$

$$w_i > 0$$

Here d_{ij} are category weights and w_i are observation weights.

Abbreviating the probabilities P_{ik} = Pr $[Y_i = k]$ allows the gradient of the log likelihood function to be expressed as

$$\frac{\partial \ln L}{\partial \hat{\mathbf{a}}_k} = \sum_{i=1}^{n} w_i (d_{ik} - P_{ik}) \mathbf{x}_{ik}$$

Similarly, the Hessian of the log likelihood has form

$$\frac{\partial \ln L}{\partial \hat{\mathbf{a}}_k \partial \hat{\mathbf{a}}'_k} = -\sum_{i=1}^{n} w_i P_{ik} (1 - P_{ik}) \mathbf{x}_{ik} \mathbf{x}'_{ik}$$

and $$\frac{\partial \ln L}{\partial \hat{\mathbf{a}}_k \partial \hat{\mathbf{a}}'_l} = -\sum_{i=1}^{n} w_i P_{ik} P_{il} \mathbf{x}_{ik} \mathbf{x}'_{il} \qquad k \neq l$$

Parameter estimation is via Newton's method.

Appendix 2

Sitka spruce / Systematic and selection thinning growth model

$$\mathbf{P}=\begin{bmatrix} 1.0 & -129.8534 & -6.4478 \\ -6.8276E-4 & 1.0 & 7.0968E-3 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\Lambda=\text{diag}(-1.393239, -0.427852, -1.0)$$

$$\mathbf{P}^{-1}=\begin{bmatrix} 1.0973 & 142.4860 & 6.0639 \\ 7.4918E-4 & 1.0973 & -2.9566E-3 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\mathbf{C}=\begin{bmatrix} 0.8291 & 1.0813\text{E}-2 & 0.2445 \\ -3.0276E-3 & 7.8987\text{E}-2 & -4.0730E-3 \\ 0.0 & 0.0 & 0.624157 \end{bmatrix}$$

$$\mathbf{C}^{-1}=\begin{bmatrix} 1.2055 & -0.1650 & -0.4733 \\ 4.6208E-2 & 12.6539 & 6.4474E-2 \\ 0.0 & 0.0 & 1.602160 \end{bmatrix}$$

$$\mathbf{a}=\begin{bmatrix} 102.0838 \\ 1.4609 \\ 11.3990 \end{bmatrix}$$

$$t_0=-0.75$$

$$\gamma_1 = -3793.0663 \qquad \gamma_2 = 0.02019244$$

Box 1: Sitka spruce (thinned) growth component coefficients.

Sitka spruce / un-thinned growth model

$$\mathbf{P}=\begin{bmatrix} 1.0 & 127.5287 & -16.4235 \\ 6.1473E-4 & 1.0 & -3.2033E-2 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\Lambda=\mathrm{diag}(-0.724367,\ -0.285728,\ -1.0)$$

$$\mathbf{P}^{-1}=\begin{bmatrix} 1.0851 & -138.3768 & 13.3880 \\ -6.6702E-4 & 1.0851 & 2.3803E-2 \\ 0.0 & 0.0 & 1.0 \end{bmatrix}$$

$$\mathbf{C}=\begin{bmatrix} 0.8226 & -7.0468E-3 & 0.4482 \\ -1.2591E-2 & -0.1874 & 0.1725 \\ 0.0 & 0.0 & 0.624157 \end{bmatrix}$$

$$\mathbf{C}^{-1}=\begin{bmatrix} 1.2149 & -4.5680E-2 & -0.8598 \\ -8.1621E-2 & -5.3327 & 1.5325 \\ 0.0 & 0.0 & 1.602160 \end{bmatrix}$$

$$\mathbf{a}=\begin{bmatrix} 248.4541 \\ 0.7077 \\ 11.3990 \end{bmatrix}$$

$$t_0=-0.75$$

$$\gamma_1=-2344.7663 \qquad \gamma_2=0.02390764$$

Box 2: Sitka spruce (un-thinned) growth component coefficients.

Sitka spruce / thinned assortment equation

Category 1 (>= 20 cm SED); parameters $\hat{\mathbf{a}}_1$

INTERCEPT	10.3553
$1/(D_0/100)$	-2.4317
$D_0 * D_0*H/10000$	-0.3571
$D_0 * D_0 / H$	0.0049
S	0.0929
$D_1 - D_0$	-4.5091
$(D_1 - D_0)$^2	0.4267
$(D_1 - D_0)/(H/10)$	4.6278
$1/((D_1 - D_0)*H)$	-7.1732
$H/(100*(D_1 - D_0)$^2)	0.1023
$(B_0 - B_1)$	0.3142
$(N_0 - N_1) / 1000$	-8.1327

Category 2 ([14, 20) cm SED); parameters $\hat{\mathbf{a}}_2$

INTERCEPT	5.1717
$1/(D_0/100)$	-0.6686
$D_0 * D_0 / 100$	-0.0720
$H/\sqrt{D_0}$	-0.2968
$D_0 * D_0*H/10000$	-0.0945
$1.0 / ((D_0 / 100)*(H/10))$	-0.2310
S	0.0520
$D_1 - D_0$	-0.5238
$(D_1 - D_0)$ ^ 2	0.1102
$(D_1 - D_0)/(H/10)$	-0.3643
$B_0 - B_1$	0.1351
$(N_0 - N_1) / 1000$	-2.0790

Box 3: Thinned Sitka spruce assortment equation.

Sitka spruce / un-thinned assortment equation

Category 1 (>= 20 cm SED); parameters $\hat{\mathbf{a}}_1$

INTERCEPT	3.2721
$1/(D_0/100)$	0.6276
D_0 / H	2.9686
$H/\sqrt{D_0}$	-0.8298
$1/((D_0/100)*(H/10))$	-1.9552
S	-0.0111
$10*D_0/N_0$	-1.1417
D_0*D_0*H/N_0	0.0128
B_0	-0.0052

Category 2 ([14, 20) cm SED); parameters $\hat{\mathbf{a}}_2$

INTERCEPT	-27.6178
D_0	1.9530
$1/(D_0/100)$	2.2596
$H/\sqrt{D_0}$	-1.3761
$D_0*D_0*H/10000$	1.1733
$1/((D_0/100)*(H/10))$	-0.9512
S	0.0090
$10*D_0/N_0$	-0.1599
$100*B_0/N_0$	-4.6630

Box 4: Un-thinned Sitka spruce assortment equation.