The Method for Deriving Theory-Based Base-Age Invariant Polymorphic Site Equations with Variable Asymptotes and other Inventory Projection Models

Plantation Management Research Cooperative
Warnell School of Forest Resources
University of Georgia

PMRC Technical Report 1999 – 4

Prepared by

C.J. CIESZEWSKI and R.L. BAILEY

March 3, 2000

Abstract

Biologically realistic site models require the ability to concurrently express variable asymptotes and polymorphism in curve shapes. Moreover, it is only logical and rational to require these models be invariant to changes in the index or base age. This manuscript explains the Generalized Algebraic Difference Approach that can be used effectively to derive truly base-age invariant difference equations capable of describing concurrent polymorphism and variable asymptotes. This new generic methodology for derivation of even the most complex dynamic equations is mathematically sound. The equations derived with it can be extremely flexible and may generate intricate patterns of concurrent polymorphism and variable asymptotes. This methodology is relevant to all situations in which the dependent variable is a function of an unobservable variable, and the models can be implicitly defined by their initial conditions. It is equally useful for derivation of new equations and for improvement of existing base-age specific equations.

Keywords: site index; base-age invariance; mixed effects; inventory updates; growth and yield.

Dynamic Equation Modeling

Background

The earliest efforts in growth and yield modeling concentrated on two-dimensional relationships such as for example height over age. Both, hand drawn curves and the earliest equations that were capable of consistently generating more intricate shapes approximated two-dimensional relations (e.g., Hosfeld 1822). These site models, at times, were developed separately for different sites or even individually for different stands. Until today many users of site models prefer to consider them as two-dimensional relationships and rarely discuss them in a context of three-dimensional systems; sometimes sites are denoted as discrete quality or productivity classes, e.g., **A**, **B**, ..., or **I**, **II**, ...

Historically, site models were presented as graphs or tables representing a four-variable height prediction system for a discrete collection of sites, or stands (e.g., USDA 1929). The variables were reference-height (discrete); age of reference-height (discrete or continuous); prediction age (continuous); and prediction height (continuous). However, these systems describe three-dimensional relationships between height, age and growth intensity, although, the growth intensity is seemingly defined by two implicit variables.

For some applications, guide curves were proportionally or otherwise (e.g., Osborne and Schumacher 1935) adjusted for individual stands by a simple means of multiplication. For example, in an anamorphic system the curve closest to an observation was multiplied by a ratio of the observation and the corresponding height on the curve at the same age. The new curve generated in this way would pass through the known height-age pair. Algebraically adjusting a single base model to specific situations or stands by scaling, improves the parsimony, consistency and utility of the system over the other approach of multiple-models developed individually for each stand. This approach reduces the number of models involved in the prediction system and, in the analysis phase, allows data from different stands to be combined in a complementary system. It extends the discrete reference-variable to a continuous reference-variable through the simple multiplication and is therefore more functional.

Newer approaches to site-dependent modeling almost exclusively involve three-dimensional functions. Usually, the models describe relationships between a response variable, time or age, and at least one more variable representing intensity of the modeled processes (e.g., García 1983). The response variable can be height, diameter, basal area, volume, number of trees, weight or any other measurable quantity. The variable representing intensity of the processes is usually for convenience expressed as an implicit measure in a form of a sample observation of the response variable. For example, a height model can use site index (S) at a given base-age (A_b). An early algebraic inclusion of S into simple anamorphic equations was followed by increasing equation complexities necessary to describe numerous desirable characteristics. Some of these characteristics include

1) curves through the origin;

2) polymorphism;

3) variable asymptotes;

4) equality of S and predicted height at base-age; and

5) theoretical justifications and interpretations.

Bailey and Clutter (1974) introduced the concept of **base-age invariance** in which a dynamic equation can compute predictions directly from any age-height pair without compromising consistency of the predictions. Dynamic equations can be viewed as four-variable relationships that are continuous over all variables. Yet, dynamic equations simply describe three-dimensional surfaces much the same as the other fixed-base-age site equations. In these equations, one of the dimensions uses two implicit variables.

The predictions of dynamic equations are unaffected by arbitrary changes in base-age. Bailey and Clutter (1974) applied a technique that has become know as the algebraic difference approach (ADA). Site equations derived with this approach are mathematically sound (i.e., it cannot lead to $1 = 0$) and they always compute consistent numbers.

We present here the Generalized Algebraic Difference Approach (GADA) — a new generic methodology for the derivation of very flexible dynamic equations that are truly base-age invariant, polymorphic, have variable asymptotes and other desired properties, such as a theoretical basis and S equal height at base-age which could not all be derived with the traditional ADA. We will show how to systematically derive these equations, as opposed to creating them in *ad hoc* ways, while neither compromising their mathematical soundness nor sacrificing any logical relationships among the equation's variables. The derivation proposed here is suitable for all kinds of growth, yield, and decline or oscillation models that could be considered for any pooled cross-sectional and longitudinal data with one unobservable variable and mixed effects modeling.

The Base-Age Invariance of Equations

The Concise Oxford Dictionary of Mathematics (Clapham 1996) defines **invariant** as *"A property or quantity that is not changed by one or more specified operations or transformations."* Other dictionaries contain consistently similar definitions. The notion of invariance in mathematics has an unequivocal meaning consistent with its dictionary and encyclopedic (e.g., Gellert *et al.* 1977) definitions; it means that the elements of interest called "invariants" in a given system will remain unaffected while other elements in the system are varied. In base-age invariant equations the **"invariants"** are the computed heights and, as a result, the shapes of the height over age curves. The varied elements are the base-ages and the reference heights.

Site curves are base-age invariant if and only if they are unequivocally unaffected by all choices of base-ages—rather than slightly affected, unaffected within a certain range, or similar to unaffected. This means that, in a dynamic equation, any arbitrary age-height pair on a curve must define the very same curve, not merely "similar curves". The base-age-invariant equations also have the path invariance property (Clutter *et al.* 1983). This means that one-step predictions or yearly or decadal iterations will all predict the same values at a given final age. When curves generated by using different base-ages are not positively identical, the equation is simply not base-age invariant.

True base-age invariant equations constitute initial condition difference equations, or dynamic equations, that are the most modern and the most advanced forms of integral basic site equations—predicting height as a function of age and site—used in forest biometrics. They represent a continuous four-variable prediction system directly interpreting three-dimensional surfaces without explicit knowledge of the third dimension, which depends on an unobservable variable. They are applicable to integral modeling, dynamic equation modeling, and to state space modeling (Garcia 1994) as well as to descriptions of complex infinite-dimensional processes in periodic or yearly iterations (e.g., Cieszewski and Bella 1993). Unlike yearly difference or differential equations, they can be used directly for forward and backward estimations.

Considering the utility and popularity of models for the dynamics of stands and forests, the literature on deriving dynamic equations is rather limited. Such derivations have been applied in different contexts. In a broad classification, the approaches to these derivations can be listed as:

- initial condition solutions to differential equations, e.g., Lenhart (1968, 1972);
- the algebraic difference approach of Bailey and Clutter (1974);
- equating sub-defined ratios of base equations, e.g., Amaro *et al.* (1998); and
- initial-condition site index substitution in expanded dynamic equations, e.g., Cieszewski and Bella (1989).

We discuss further only the algebraic difference approach.

The Algebraic Difference Approach

The dynamic equations depend on their entangled self-definition, i.e., the inverse function of their underlying base equation that is entangled into the base equation. They are extremely sensitive to any *ad hoc* algebraic operations that are otherwise harmless with all explicit equations. This sensitivity goes far beyond what "usual" operations with explicit equation based common sense may dictate. For example, these dynamic equations cannot be added by sides or altered in other ways by adding arbitrary exponentiation or multiplication within the existing structures. They cannot be created by assuming or relaxing any implicit assumptions or in other *ad hoc* ways. All of the above would be perfectly admissible with explicit equations but not with these implicitly defined dynamic equations.

The Algebraic Difference Approach (ADA) is similar to the standard mathematical procedure applied in calculus for boundary solutions to differential equations. It consists of replacing an arbitrary parameter in a function Y of t with its solution using specific values of Y_0 and t_0 instead of Y and t.

Writing the base function Y as a function of t and n parameters $\rho_1 \dots \rho_n$ as:

$$Y(t) = f(t, \rho_1 \dots \rho_{n-1}, \rho_n), \tag{1}$$

its solution for an arbitrary parameter ρ_n is a function of two independent variables Y and t and the $n-1$ remaining parameters. In this solution, Y and t are independent variables and therefore can be assigned arbitrary values Y_0 and t_0:

$$\rho_n = u(t, Y, \rho_1 \dots \rho_{n-1}) = u(t_0, Y_0, \rho_1 \dots \rho_{n-1}) \tag{2}$$

where Y_0 denotes a given value of Y for an arbitrary t_0.

The above solution can be used in place of ρ_n in the base function to define a new dynamic function of time t, an arbitrary time t_0, a given function value Y_0 at t_0 and the remaining $n-1$ parameters:

$$Y(t, t_0, Y_0) = w(t, t_0, Y_0, \rho_1 \dots \rho_{n-1}) \tag{3}$$

Since the above function is undefined without the arbitrary values of t_0 and Y_0, these values are called the initial conditions of this function. With t_0 and Y_0 assuming any values, the equation represents a dynamic equation. Using this approach, a site index equation may be developed that produces a site index curve unchanging under all choices of base-age. Hence, the term base-age invariant equation.

The ADA was introduced to forestry literature using a logarithmic transformation of the Schumacher (1939) equation, i.e.,

$$\ln Y(t) = \alpha - \beta/t \tag{4}$$

where Y was height but it could be any variable of interest (e.g., volume, density, basal area, etc.), and t was defined as a simple exponential function of age (Age^c where c was a parameter). The solution for the proportionality constant of this function is

$$\alpha = \ln Y_0 + \beta/t_0.$$

Using ADA, the anamorphic equation with variable asymptotes is

$$\ln Y(t, t_0, Y_0) = \ln Y_0 + \beta(1/t_0 - 1/t). \tag{5}$$

A polymorphic equation (with a single asymptote) may be obtained by applying ADA to the β parameter of the equation. The solution of eq. (4) for the slope parameter is

$$\beta = t_0(\alpha - \ln Y_0)$$

so the polymorphic dynamic equation is

$$\ln Y(t, t_0, Y_0) = \alpha + (\ln Y_0 - \alpha)t_0/t \tag{6}$$

The above equations derived by ADA represent true equalities. They are consistent in predictions and any point on a curve will always unequivocally define the very same curve. The equations also have the path invariance property described earlier. In short, the curves are indifferent under choices of base-ages; they are base-age invariant.

The algebraic difference approach is, in principle, similar to the method used for solutions of differential equations with initial conditions, but as applied to base equations it moves outside of integration theory. For example, a general solution to the differential equation (4) of McDill and Amateis (1992):

$$\frac{dY}{dt} = \alpha\frac{Y}{t}\left(1 - \frac{Y}{M}\right)$$

should be presented with an intercept defined by the integration constant C:

$$Y(t) = \frac{M}{1 + \beta/t^\alpha} + C$$

However, defining the initial conditions based on C, as integration theory would suggest, is unreasonable in those cases where biological interpretation requires that $Y(0) = 0$ (i.e., $C = 0$). Thus, McDill and Amateis (1992) applied the algebraic difference approach to the parameter β and using its initial condition solution:

$$\beta = {t_0}^\alpha\left(\frac{M}{Y_0} - 1\right)$$

they have derived their polymorphic dynamic equation [5][1], i.e.,

$$Y(t) = \frac{M}{1 + \left(\frac{M}{Y_0} - 1\right)\left(\frac{t_0}{t}\right)^\alpha}$$

The ADA has been applied successfully in various modeling contexts by Bégin and Schütz (1994); Borders *et al.* (1984 and 1988); Cao *et al.* (1993, 1997); Clutter *et al.* (1983); Clutter *et al.* (1984);

[1] The reference [5] is for the original publication not this text.

DuPlat and Tran-Ha (1986); Lappi and Bailey (1988); McDill and Amateis (1992); Ramirez *et al.* (1987); and others cited therein. These workers have used the ADA and the dynamic equations for modeling growth and yield of height, diameter and basal area, as well as tree survival in forest populations.

The New Methodology

Theoretical Foundation and Symbolic Definition

To facilitate the Generalized Algebraic Difference Approach (GADA) formulation we identify a theoretical variable labeled the growth intensity factor $\mathcal{X}$ and define it to be the quantification of those particular growth dynamics that are uniquely associated with a site and individual characteristics of growth or survival capabilities.

$\mathcal{X}$ is used consistently in all equation formulations to describe the rules of changes in curve shapes across different sites. It can be either a variable or a function of any number of variables. Such variables can include climate, water availability, organic soil depth, leaf area or rates of photosynthesis, measure of ozone and genetic components. $\mathcal{X}$ is continuous, monotonic and relevant to the modeled dynamics; it can describe the relative rates of change in terms of direct functional relationship. We can assume that, for example, small values for $\mathcal{X}$ represent low growth intensity and high values for $\mathcal{X}$ represent high growth intensity. Not being practically obtainable, $\mathcal{X}$ is eventually replaced with the initial conditions that are measurable so that the equation can be operationally useful. However, this happens only after the equation is explicitly formulated in a satisfactory way when it already contains all the desired properties of a site equation, such as, polymorphism and variable asymptotes.

The first step in the GADA is to select a base equation and to identify in it any desired number of site-specific parameters. Then, define explicitly how the site-specific parameters change across different sites by replacing them with explicit functions of $\mathcal{X}$ and new parameters. In this way the initially selected two-dimensional base equation expands into an explicit three-dimensional site equation describing both cross-sectional and longitudinal changes with two independent variables t and $\mathcal{X}$. In the final step a solution for $\mathcal{X}$ replaces all $\mathcal{X}$s with implicit definitions using the equation's initial conditions t_0 and Y_0.

Symbolically, the base equation is

$$Y(t) = f(t, \rho_1 \dots \rho_{n-1}, \rho_n) \tag{7}$$

where $\rho_1 \dots \rho_n$ are the equation parameters.

If in the base equation (7) a given site-specific parameter ρ_i is defined as a function g_i of $\mathcal{X}$ and any number of j new parameters, viz., $\rho_i \equiv g_i\big(\mathcal{X}, \rho_{i_1} \dots \rho_{i_j}\big)$, the base equation (7) with multiple site-specific parameters is changed to the explicit three-dimensional site equation with two independent variables t and $\mathcal{X}$:

$$Y(t, \mathcal{X}) = f\Big(t, \rho_1 \dots \rho_{m-1}, g_m(\mathcal{X}, \rho_{m_1} \dots \rho_{m_k}) \dots g_n(\mathcal{X}, \rho_{n_1} \dots \rho_{n_l})\Big) \tag{8}$$

where $Y(t, \mathcal{X})$ is a function of t, $\mathcal{X}$, and $m + k + l - 1$ parameters.

If eq. (8) can be solved for $\mathcal{X}$, the RHS of this solution, with initial condition values for t and Y, i.e.,

$$\mathcal{X} = u(t, Y, \rho_1 \dots \rho_{n_l}) = u(t_0, Y_0, \rho_1 \dots \rho_{n_l}), \tag{9}$$

can be substituted in eq. (8) in place of $\mathcal{X}$ so the dynamic equation

$$Y(t, t_0, Y_0) = f\Big(t, \rho_1 \dots \rho_m, u(t_0, Y_0, \rho_1 \dots \rho_{n_l})\Big),$$

after reformulation and elimination of redundant parameters, becomes the dynamic equation with an implicitly defined initial condition:

$$Y(t, t_0, Y_0) = f(t, t_0, Y_0, \rho_1 \dots \rho_w) \tag{10}$$

where

$$n - 1 \le w \le m + k + \dots + l - 1 \qquad \text{and} \quad k \lesseqqgtr \dots \lesseqqgtr l. \tag{11}$$

The result in eq. (11) means that equation (10) has a smaller or equal number of parameters than the equation (8).

Practical applications of the GADA involve different levels of complexity and difficulty in equation derivations. We classify the equations as simple or complex depending on whether the solutions involved are based on just a reformulation of an equation (simple) or on finding its roots (complex).

Specific Cases

Simple Equations

In the simplest applications the advantage of introducing $\mathcal{X}$ is not immediately obvious. For example, to replicate the Bailey and Clutter (1974) derivation based on two equations with the GADA, we write eq. (4) in two ways:

$$\ln Y(t, \mathcal{X}) = \mathcal{X} - \beta_a/t \tag{12}$$

and

$$\ln Y(t, \mathcal{X}) = \alpha_p - \mathcal{X}/t \tag{13}$$

where β_a is the slope parameter of the anamorphic equation and α_p is the asymptote parameter of the polymorphic equation. Applying the ADA with respect to $\mathcal{X}$ in either of these two equations completes the process and concludes this application of the GADA. The two dynamic equations in Bailey and Clutter (1974) may be derived in this way.

The greatest advantage of introducing $\mathcal{X}$ manifests itself when more than one simultaneous site-specific parameter is necessary to adequately describe changes in curve shapes across different sites. For example, in a simple assumption of concurrent polymorphism with varying asymptotes both α and β in

$$\ln Y(t) = \alpha - \beta/t$$

could be dependent on $\mathcal{X}$ while $\mathcal{X}$ could define the limiting size, i.e.,

$$\ln Y(t, \mathcal{X}) = \mathcal{X} + \beta\mathcal{X}/t \tag{14}$$

The solution for $\mathcal{X}$ would then be

$$\mathcal{X} = \frac{\ln Y}{1 - \beta/t} = \frac{\ln Y_0}{1 - \beta/t_0}$$

and applying the GADA to eq. (14) with respect to $\mathcal{X}$ would result in a dynamic equation based on Schumacher's equation that provides polymorphic base-age invariant curves with variable asymptotes:

$$\ln Y(t, t_0, Y_0) = \ln Y_0 \frac{t_0(t - \beta)}{t(t_0 - \beta)} \tag{15}$$

The assignment of $\mathcal{X}$ to α means that given an objective measure of growth intensity the upper production limit would be increasing with increasing innate growth potential. This would result in variable asymptotes. The assignment of $\mathcal{X}$ to β means, in simple terms, that the shapes of curves change with changing growth intensity which defines a polymorphic equation. Clearly, both variable asymptotes and polymorphism occur if $\mathcal{X}$ affects both α and β.

Alternatively, the objective could be a single equation that concurrently expresses 1) a similar polymorphism to that of the Bailey and Clutter (1974) polymorphic equation; and 2) similar asymptotic properties to those of the Bailey and Clutter (1974) anamorphic equation. The advantage of introducing $\mathcal{X}$ becomes most evident here as this objective is accomplished simply by adding eqs. (12) and (13) by sides, i.e.,

$$2 \ln Y(t, \mathcal{X}) = (\mathcal{X} - \beta_a/t) + (\alpha_p - \mathcal{X}/t) \tag{16}$$

Thus, the solution

$$\mathcal{X} = \frac{t(\ln Y - \alpha'_p) + \beta'_a}{t - 1} = \frac{t_0(\ln Y_0 - \alpha'_p) + \beta'_a}{t_0 - 1}$$

substituted for $\mathcal{X}$ in eq. (16) produces a single dynamic equation exhibiting concurrently both of the desired properties:

$$\ln Y(t, t_0, Y_0) = \alpha'_p - \frac{\beta'_a}{t} + \frac{(t-1)t_0}{(t_0 - 1)t}\left(\ln Y_0 - \alpha'_p + \frac{\beta'_a}{t_0}\right) \tag{17}$$

The ability to combine the properties of two different dynamic equations into one dynamic equation by adding their explicit forms by sides is a unique advantage of the new methodology. However, this must be done in the explicit stage with the base equations. If the sides of two dynamic equations are added directly the result is a degenerated[2] relationship that has neither the property of base-age invariance nor that of equality. Adding equations by sides is improper when applied directly to dynamic equations. It is admissible only when applied to base equations before entangling the equations with initial conditions. Although not discernible with the basic algebraic difference approach the above dynamic equation becomes clearly apparent with the Generalized Algebraic Difference Approach.

Complex Equations

We label dynamic equations as simple when they can be derived through direct reformulation as shown above. All dynamic equations that require in their derivations the roots of an equation in order to determine $\mathcal{X}$ we label complex because the type of solutions required may constitute a considerable barrier in practical applications.

Even with few parameters, an equation can be complex with solutions that involve roots. Examples occur in formulations involving quadratic relationships or combinations of direct and inverse proportionality. Such a relationship could be between an equation characteristic, e.g., a variation in asymptotes or polymorphism, and a growth intensity measure $\mathcal{X}$. For example, the derivation of a complex dynamic equation could follow from a theory that asymptotes are exponentially proportional to growth intensity, and that polymorphism is inversely proportional to growth intensity, i.e.,

[2] "improper", "inadmissible", "degenerate", and so on, are operations and/or formulations that can lead to $1 = 0$, after Simmons (1972) (p. 160, l. 17 and bottom of the page) who writes "...This terminology follows a time-honored tradition in mathematics, according to which situations that elude simple analysis are dismissed by such pejorative terms as "improper," "inadmissible," "degenerate", "irregular," and so on...."

$$\ln Y(t, \mathcal{X}) = \alpha \mathcal{X} - \frac{\beta/\mathcal{X}}{t} \tag{18}$$

For this base equation, the solution for $\mathcal{X}$ involves finding roots of a quadratic equation and a selection of the most appropriate root to entangle into the dynamic equation. The selection of the most appropriate expression for $\mathcal{X}$ may depend on the equation parameters that in turn depend on the data and the domain of the applicable ages. The solution for $\mathcal{X}$ in eq. (18) is

$$\mathcal{X} = \begin{cases} \begin{cases} 0.5(\ln y + \mathcal{R})/\alpha = 0.5(\ln Y_0 + \mathcal{R}_0)/\alpha \\ \text{or:} \\ 0.5(\ln y - \mathcal{R})/\alpha = 0.5(\ln Y_0 - \mathcal{R}_0)/\alpha \end{cases} \\ \text{where:} \\ \begin{cases} \mathcal{R} = \sqrt{(\ln Y)^2 + 4\,\alpha\,\beta/t} \\ \text{and:} \\ \mathcal{R}_0 = \sqrt{(\ln Y_0)^2 + 4\,\alpha\,\beta/t_0} \end{cases} \end{cases}$$

Selecting the root more likely to be real (as opposed to complex) and positive, i.e., the one involving addition rather than subtraction of the square-root, with the usual initial conditions and substituting it into eq. (18) results in the following dynamic equation:

$$\ln Y(t, t_0, Y_0) = \frac{\ln Y_0 + \mathcal{R}_0}{2} - \frac{2\gamma/t}{\ln Y_0 + \mathcal{R}_0} \tag{19}$$

where $\gamma = \alpha\beta$.

Another situation requiring a root-finding solution arises when the cross-sectional changes are described by polynomial functions of $\mathcal{X}$.

The pursuit of a best equation form may become a tedious procedure depending on many factors including the data analysis. For each explicit or base equation, several possible approaches may be used to derive the implicit dynamic equation. However, at any time a new implicitly defined equation is considered, the formulation of proper relationships in the explicit equation should be completed prior to the entangling of implicit solutions. For this work, a good understanding of the explicit equation's mathematical structure and the biological expectations of growth differences over different sites give the modeler a distinct advantage. However, absent an understanding or knowledge of the expected growth relationships, curve shapes desired, or the functional changes wanted, one may explore a formulation of generic relationships such as we discuss in subsequent sections. After parameter estimation with such a model, one may then exercise hypothesis testing based on estimates of model parameters and their error structures to resolve questions about biological behavior of the system.

Multiple and Stepwise Regression Equations

The methodology we advocate defines a rigorous mathematical procedure facilitating the derivation of equations with implicitly defined initial conditions from explicit theoretical bases relating to biological, geometric or algebraic theories. This methodology emphasizes the role of the modeler in formulating the hypothesis upon which the equations are built prior to their final restructuring into dynamic equations. The equations are formed by the modeler rather than by default via statistical analysis. However, the "generic equations" discussed later are intended to provide an excessive amount of flexibility in anticipation that statistics of fit will determine the final forms of the dynamic equations.

It may be necessary in practice for statistics rather than the modeler to determine forms of final equations. This may apply not only to stepwise and permutational regressions but also to any other type of linear or nonlinear regression analysis or model fitting in which the criteria for model selection depend on residual analysis or statistical results. Given such a situation it may seem that the GADA is antithetical to regression theory. We believe otherwise.

The methodology presented here can be used to improve existing regression equations even if they are produced by step-wise regressions. With stepwise regression in particular, the GADA may have considerable value in equation improvement efforts. Consider, for example, the following four-parameter equation based on stepwise-regression:

$$Y(t, S) = \alpha\sqrt{t} + \beta t^2 \ln^{32} t - \gamma\frac{t^{5/2}}{\ln t} + S\delta\sqrt{t} \tag{20}$$

with solution

$$S(t, Y) = -\frac{\alpha\sqrt{t_0}\,\ln t_0 + t_0\beta t_0^2 \ln^{33} - t_0^{5/2} - \gamma Y_0\,\ln t_0}{\delta\,\ln t_0\sqrt{t_0}} \tag{21}$$

giving rise to the two-parameter dynamic equation

$$Y(t, t_0, Y_0) = \beta\left(t^2 \ln^{32} t - t_0^{3/2} t_0 \ln^{32} \sqrt{t}\right) + \gamma\left(\frac{t_0^2\sqrt{t}}{\ln t_0} - \frac{t^{5/2}}{\ln t}\right) + Y_0\sqrt{\frac{t}{t_0}} \tag{22}$$

The GADA methodology has been used to convert the model into a dynamic relationship that:

1. Generates identical curves as those produced by eq. (20)

2. Predicts heights at base-age equal to site indexes

3. Can compute site index and height from the same equation

4. Can use heights and ages directly instead of fixed-base-age site indexes and

5. Can be easily fitted and applied with use of any base-age.

All these improved properties accrue with a reduction by half in the number of parameters and no contradiction to regression theory or practice.

Generic Equations

In this section, we present the most advanced category of dynamic base equations. They may be considered the epitome of equation-based modeling with dynamic equations as discussed herein. Generic equations are formulated in the absence of explicit expectations about the final model form. A modeler may want to cover a wide range of possible equations during a single analysis to save time and make equation selection more efficient. Schnute (1981) discusses an excellent example of such a practice. Generic equations have been considered here as a separate category because of their potentially large number of parameters and complicated appearance resulting from either simple or complex derivations. Development of generic equations should be considered with caution because it can easily lead to over-parameterization and model instability as well as difficulties with parameter estimation.

An example founded on the Schumacher base equation might represent a lack of strong commitment to either the asymptote or the shape parameter being the only or most predominant expression of growth intensity. In other words, it may be appropriate to derive an equation that might, but does not have to, have asymptotes affected by site factors and might, but does not have to, have curve shapes varying across different sites. In addition, these effects could occur in flexible proportions. A simple base equation that satisfies such requirements is a generalization of eq. (16):

$$\ln Y(t, \mathcal{X}) = (\alpha + \alpha' \mathcal{X}) - (\beta + \beta' \mathcal{X})/t \tag{23}$$

where α' and β' are the weighting parameters. This generic form of the explicit three-dimensional equation can easily examine what proportions of eqs. (12) and (13), or eqs. (4) and (14), are best blends for any given data. This equation also allows one to examine if eq. (14) should indeed be directly proportional to $\mathcal{X}$ or if it should be only linearly (or partially) proportional to $\mathcal{X}$. To illustrate these three alternative hypothesis, eq. (16) can be written as a weighted sum of eq. (12) and eq. (13). This produces a reparameterized version of eq. (23):

$$\ln Y(t, \mathcal{X}) = \alpha' (\mathcal{X} - \beta_a/t) + \beta' (\alpha_p - \mathcal{X}/t)$$

(where: $\alpha_p = \alpha/\beta'$ and $\beta_a = \beta/\alpha'$); it can be written as a weighted sum of eq. (4) and eq. (14), which is also equivalent to eq. (23):

$$\ln Y(t, \mathcal{X}) = \alpha' \mathcal{X} (\alpha - \beta/t) + \beta' (\alpha - \beta/t)$$

and it can be written as a linear generalization of eq. (14):

$$\ln Y(t, \mathcal{X}) = (\alpha' \mathcal{X} + \beta') (\alpha - \beta/t)$$

where: α' and β' are the weights of the anamorphic and polymorphic forms and $\alpha_p = \alpha/\beta'$ and $\beta_a = \beta/\alpha'$ and both $\alpha' \neq 0$ and $\beta' \neq 0$. The solution for $\mathcal{X}$ in eq. (23) is

$$\mathcal{X} = \frac{\ln Y - \alpha - \beta/t}{\alpha' - \beta'/t} = \frac{\ln Y_0 - \alpha - \beta/t_0}{\alpha' - \beta'/t_0}$$

and after applying the GADA to eq. (23) and using this solution, the resulting simple generalized dynamic equation based on eq. (4) has the following form:

$$\ln Y(t, t_0, Y_0) = \alpha - \frac{\beta}{t} + \frac{\alpha' - \beta'/t}{\alpha' - \beta'/t_0} \left(\ln Y_0 - \alpha + \frac{\beta}{t_0} \right)$$

However, this equation is clearly over-parameterized to the extent of being undefinable. This can be rectified by combining the parameters α' and β' into one parameter. Depending on which of the two parameters in eq. (23) is more likely to be equal to zero the corresponding dynamic equation could have one of the two forms:

$$\ln Y(t, t_0, Y_0) = \begin{cases} \alpha - \beta/t + (\ln Y_0 - \alpha + \beta/t_0)\,(1 - \gamma/t)\,/\,(1 - \gamma/t_0) \\ \text{for expected: } \alpha' \neq 0 \\ \text{or:} \\ \alpha - \beta/t + (\ln Y_0 - \alpha + \beta/t_0)\,(\delta - 1/t)\,(\delta - 1/t_0) \\ \text{for expected: } \beta' \neq 0 \end{cases} \tag{24}$$

where: $\gamma = \beta'/\alpha'$ and $\delta = \alpha'/\beta'$ and at least one of the two parameters must be different from zero. If both $\alpha' = 0$ and $\beta' = 0$ there is no site equation defined by eq. (23) but rather a simple two-dimensional single-line equation that does not involve the concept of site index or base-age invariance. That is to say, the data either represent a single site or a series of different sites containing excessive amounts of crossing or noise rendering unique identification of separate sites impossible.

Hypothesis testing on equation (24) may be carried on by means of simple tests of significance for different model parameters. Some potential outcomes from such tests could be:

- $\beta' = 0$: The equation is anamorphic with variable asymptotes.

- $\alpha' = 0$: The equation is polymorphic with a single asymptote.
- $\beta' \neq 0$ and $\alpha' \neq 0$: The equation is polymorphic and has variable asymptotes.
- $|\alpha'| \ll |\beta'|$: The equation exhibits relatively strong polymorphism.
- $|\alpha'| \gg |\beta'|$: The equation exhibits relatively strong identification of variable asymptotes.

An example of a complex generic equation can be developed from a generalization of eq. (18):

$$\ln Y(t, \mathcal{X}) = \alpha + \alpha' \mathcal{X} - \frac{\beta + \beta'/\mathcal{X}}{t} \tag{25}$$

The solution involves solving a quadratic equation in $\mathcal{X}$. Since there are two roots, careful consideration must be given to which is most appropriate in the final equation. The selection may depend on the model parameters, which in turn depend on the data and the domain of the applicable ages. In the above example the root most likely to be real and positive, and therefore more likely to be useful is

$$\mathcal{X} = \begin{cases} 0.5\,(\mathcal{R}_0 - \alpha)\,/\alpha' \\ \text{where:} \\ \mathcal{R}_0 = \beta/t_0 + \ln Y_0 + \sqrt{(\ln Y_0 - \alpha + \beta/t_0)^2 + 4\,\gamma/t_0} \end{cases}$$

Substituting this root for $\mathcal{X}$ (eq. (25)) results in a generalization of eq. (19), i.e., the following complex generic dynamic equation:

$$\ln Y(t, t_0, Y_0) = \frac{\mathcal{R}_0 + \alpha}{2} + \frac{2\,\gamma/t}{\mathcal{R}_0 - \alpha} - \frac{\beta}{t} \tag{26}$$

where: $\gamma = \alpha'\beta'$.

Properties of the Approach

Parsimony

The Generalized Algebraic Difference Approach is more parsimonious than most traditional approaches to site equation derivations or formulations and can derive more complex equations than the traditional Algebraic Difference Approach. In terms of the potential for final equation flexibility, it exceeds the capabilities of fixed-base-age modeling approaches. This new approach can, in various cases, produce equations that are more flexible and have fewer parameters than the corresponding to them fixed-base-age equations. An example is in eq. (22).

Our contention that the Generalized Algebraic Difference Approach is more parsimonious than the fixed-base-age approach is justified by three points.

1. **The GADA does not require any new parameters** in addition to the ones existing in the explicit- or fixed-base-age site equation to which it is applied, which is evident from the methodology definition symbolized by equations (8) to (10).

2. **The conclusion in eq. (11)** applied to derivations of dynamic equations from fixed-base-age site index equations demonstrates unequivocally that the final dynamic equation (10) has a smaller or equal number of parameters than the initial fixed-base-age equations used in eq. (8).

3. **The lack of dimensionality or range definition on $\mathcal{X}$,** assures that any multi-parameter expression involving the unobservable, multidimensional variable $\mathcal{X}$ will always be reparameterized into the most parsimonious form. For example, the three parameter relationships $\alpha\mathcal{X}^\gamma$

and $\beta\mathcal{X}^\gamma$ are automatically equivalent to the one parameter relationships $\alpha'\mathcal{X}'$ and $\mathcal{X}'$ or $\mathcal{X}'$ and $\beta'\mathcal{X}'$.

Clearly, if the approach never uses more parameters **(1.)** but sometimes uses fewer parameters **(2.** and **3.)** than another approach then it is, in general, a more parsimonious approach.

Point (1.) is based on the fact that the only step in the GADA that adds parameters is the formulation of the explicit base site equation. In a special case of the GADA, the base site equation can be formulated as a fixed-base-age site index equation and still be applicable for the dynamic equation derivation. Thus, there is no disadvantage involved in this step of the GADA.

Point (2.) accounts for such situations as the four parameter fixed-base-age eq. (20) been re-derived with the GADA as the two parameter dynamic eq. (22) with increased flexibility.

Finally, point (3.) accounts for situations in which the explicit base site equation is unintentionally over-specified, a fact that cannot be easily identified with the more traditional approaches. An example can be the Schumacher (1939) equation with $\alpha \propto \alpha'\mathcal{X}^\gamma$ and $\beta \propto \beta'\mathcal{X}^\gamma$. Since X is an unobservable variable and, unlike site index, has only a theoretical meaning not intended for explicit practical use, it can be freely redefined as either $\mathcal{X}' = \alpha'\mathcal{X}^\gamma$ or $\mathcal{X}' = \beta'\mathcal{X}^\gamma$, vis.,

$$\ln Y(t, \mathcal{X}) = \alpha\mathcal{X}^\gamma + (\beta\mathcal{X}^\gamma)/t = \mathcal{X}' + \beta'\mathcal{X}'/t \equiv \mathcal{X} + \beta\mathcal{X}/t \equiv \alpha\mathcal{X} + \mathcal{X}/t \tag{27}$$

Even if the modeler does not notice this opportunity for parameter reduction in eq. (27), the derivation defined by the GADA automatically reduces the number of parameters by cancellation of terms during routine algebraic operations. Such is not as likely to happen when dealing with fixed-base-age site index equations.

Robustness

Two aspects of the GADA approach to deriving models based on dynamic equations assure a high degree of robustness in applications. First, the theoretical variable $\mathcal{X}$ has no restrictions in interpretation. Second, the unobservable variable $\mathcal{X}$ is eliminated during the derivations. Not only are the dynamic equations derived with the GADA generalizations of many functional forms of the unobservable variable as shown above, but they are also generalizations of many, at times contradictory, theories behind the model. For example, if the applied theory were based on a proportional relationship the final dynamic equation would include this proportional relationship as a special case but would not be limited to it. The same dynamic equation would consolidate many various theories as numerous special cases that include a competing theory based on a corresponding inverse-proportional relationships.

Example of Application of the GADA to Comparing Base-Age Specific Fitting Methodologies[3]

A number of authors have addressed questions related to the various options for fitting site models. Yet, various issues remain unresolved that may be examined with the aid of the GADA.

Curtis (1990[4]) compared curves generated by two fixed-base-age site index equations fitted with

[3] This section describes a part of study conducted in 1990 by the first author in collaboration with Dr. R.O. Curtis, USDA Forest Service, who inspired the investigation described here through his questioning of different fitting methodologies for site index models and who also provided the data for such analysis.

[4] Curtis, RO, 1990. Site Index Curves From Stem Analyses—Methodology Effects and a New Technique. Talk presented on Western Mensurationist Meeting, June 20-22, 1990, Bend, Oregon, USA. Results also contained in an unpublished manuscript (rev. 5/07/1990) by R.O. Curtis.

base-age specific methodologies using base-ages 50 and 100 years. Comparison of fitting methodologies should be conducted with a common mathematical expression so the effects of the methodologies are not confounded with the effects resulting from using different mathematical formulations for each method. For this reason, Curtis derived equations that are very similar from a common base equation.

Curtis' (1990) equation for base-age 100 has a form:

$$Y(t, S_{100}) = \exp\left(\frac{\ln S_{100} + \alpha(\ln t - \ln 100) + \beta(\ln t - \ln 100)^2}{1.0 + \gamma(\ln t - \ln 100) + \delta(\ln t - \ln 100)^2}\right) \tag{28}$$

with the inverse site index prediction equation:

$$S_{100}(t, Y) = \exp\left(-\alpha(\ln t - \ln 100) - \beta(\ln t - \ln 100)^2 + \ln Y(1.0 + \gamma(\ln t - \ln 100) + \delta(\ln t - \ln 100)^2)\right) \tag{29}$$

and the equation for base-age 50 has the form:

$$Y(t, S_{50}) = \exp\left(\frac{\ln S_{50} + \alpha(\ln t - \ln 50) + \beta(\ln t - \ln 50)^2}{1.0 + \gamma(\ln t - \ln 50) + \delta(\ln t - \ln 50)^2}\right) \tag{30}$$

with the inverse site index prediction equation:

$$S_{50}(t, Y) = \exp\left(-\alpha(\ln t - \ln 50) - \beta(\ln t - \ln 50)^2 + \ln Y(1.0 + \gamma(\ln t - \ln 50) + \delta(\ln t - \ln 50)^2)\right) \tag{31}$$

Based on his analysis with the above equations Curtis concluded, among other things, that:

> " ...Reversibility, relative insensitivity to choice of reference age, and the uncertainties associated with inconsistency in errors in the predictor variables used in derivation versus those used in practical applications of conventional regressions all suggest that the structural relationship is a reasonable compromise and a plausible alternative to the more commonly used regression procedures. The structural equation has great practical advantage of providing a single equation for both site index and height growth estimates."

Two notable themes of Curtis' work, which can be markedly enhanced by an application of the Generalized Algebraic Difference Approach, are:

i) comparison of base-age dependent fitting methodologies using different base-ages; and

ii) development of models using a "single" equation for height and site index predictions.

Some observations regarding the themes in Curtis' work and approaches to their investigation are in order. First, equations (28) and (30) are similar but not the same. They will not generate identical curves even if fitted to data without error. Since they have four nonlinear parameters, they are flexible enough to appear on a graph as visually similar even though algebraically and numerically each produces a different set of co-ordinates. Ironically, only dynamic equations can provide principally pure grounds with a single common equation for directly testing base-age specific fitting techniques using various base-ages.

Second, only dynamic equations can actually provide one single equation for estimating both height and site index if such a principle is accepted within an adopted statistical framework. Although equations (28) and (29) are derived from each other and are merely inverse functions of each other, they are, in fact, two separate equations. The same applies to the equations (30) and (31). The derivation of dynamic equations allows one to specifically address these principle points and thus, enhance any study similar to Curtis'.

To analyze any or all of the four equations (28) to (31) with different base-ages one can apply the Generalized Algebraic Difference Approach to derive a dynamic generalization of these equations. First, we define two variables, the site variable $\mathcal{X}$ to take the place of $\ln S$ and the measurement base-age $\mathcal{Z}$ to take the place of the constants 100 and 50. Inclusion of these and simplification produces the following generalized explicit equation that can be base-age specific:

$$\ln Y(t, \mathcal{X}, \mathcal{Z}) = \frac{\mathcal{X} + (\alpha + \beta \ln t) \ln(t/\mathcal{Z})}{1 + (\gamma + \delta \ln t) \ln(t/\mathcal{Z})} \tag{32}$$

where: Y is any applicable variable of interest such as height or volume; $\mathcal{X}$ is the GADA "universal" unobservable site variable; $\mathcal{Z}$ is a constant or a parameter equal to 100 for eq. (28) and 50 for eq. (30). The parameters are unique to this equation and the equation is presented as the logarithmic transformation for the sake of simplicity of presentation.

Just as its special cases (i.e., eqs. (28) and (30)), eq. (32) is base-age specific and cannot be directly analyzed with different base-ages. Following the GADA, the initial condition solution for $\mathcal{X}$ in eq. (32) is:

$$\mathcal{X} = \ln(t_0/\mathcal{Z}) \left(\ln({Y_0}^{\gamma} t_0^{\delta \ln Y_0 - \beta}) - \alpha \right) + \ln Y_0 \tag{33}$$

and substitution into eq. (32) with some simplifications leads to the following dynamic generalization of the four equations (28) to (31):

$$\ln Y(t, t_0, Y_0, \mathcal{Z}) = \frac{\ln(Y_0 \ (t_0/\mathcal{Z})^{\ln Y_0 (\gamma + \delta \ \ln t_0) - \beta \ \ln t_0 - \alpha} \ (t/\mathcal{Z})^{\alpha + \beta \ \ln t})}{1 + (\gamma + \delta \ \ln t) \ln(t/\mathcal{Z})}$$

or

$$Y(t, t_0, Y_0, \mathcal{Z}) = \exp \left(\frac{\ln(Y_0 \ (t_0/\mathcal{Z})^{\ln Y_0 (\gamma + \delta \ \ln t_0) - \beta \ \ln t_0 - \alpha} \ (t/\mathcal{Z})^{\alpha + \beta \ \ln t})}{1 + (\gamma + \delta \ \ln t) \ln(t/\mathcal{Z})} \right) \tag{34}$$

The dynamic equation (34) includes an infinite number of different equations with the constant (or parameter) $\mathcal{Z}$ equal to any arbitrary or estimated real number. Four special cases of this equation are the four equations (28), (29), (30) and (31).

Depending upon imposed constraints as shown below, eq. (34) simplifies to one of the following four special cases:

1) When $t_0 = \mathcal{Z} = 100$ then Y_0 is equivalent to S_{100}, Y is equivalent to predicted height, eq. (34) is equivalent to eq. (28), and it becomes:

$$Y(t, S_{100}) = \exp \left(\frac{\ln(S_{100} \ (.01t)^{\alpha + \beta \ \ln t})}{1 + (\gamma + \delta \ \ln t) \ln(.01t)} \right)$$

2) When $t = \mathcal{Z} = 100$ then Y is equivalent to S_{100}, Y_0 is equivalent to direct height measurements, eq. (34) is equivalent to eq. (29), and it simplifies to:

$$S_{100}(t_0, Y_0) = \exp \left(\ln(Y_0 \ (.01t_0)^{\ln Y_0 (\gamma + \delta \ \ln t_0) - \beta \ \ln t_0 - \alpha}) \right)$$

3) When $t_0 = \mathcal{Z} = 50$ then Y_0 is equivalent to S_{50}, Y is equivalent to predicted height, eq. (34) is equivalent to eq. (30), and it becomes:

$$Y(t, S_{50}) = \exp \left(\frac{\ln(S_{50} \ (.02t)^{\alpha + \beta \ \ln t})}{1 + (\gamma + \delta \ \ln t) \ln(.02t)} \right)$$

4) When $t = \mathcal{Z} = 50$ then Y is equivalent to S_{50}, Y_0 is equivalent to direct height measurements, eq. (34) is equivalent to eq. (31), and it simplifies to:

$$S_{50}(t_0, Y_0) = \exp\left(\ln(Y_0\ (.02t_0)^{\ln Y_0(\gamma+\delta\ \ln t_0)-\beta\ \ln t_0-\alpha})\right)$$

Equation (34) is base-age invariant with respect to its initial conditions, i.e., the internal base-age of the equation, t_0 and Y_0. However, as with any other equation, the values calculated with the given formula are subject to the values of the parameters and constants of the formula. The value of $\mathcal{Z}$ addresses differences between eq. (28) and eq. (30) as well as many other similar equations. However, the fact that eq. (34) is base-age invariant allows each of these equations to be fitted with a base-age specific regression methodology using any arbitrary base-age selection.

The equivalent of eq. (28) can be fitted using eq. (34) as a base-age specific site index equation. For a base-age of 100 years, $t_0 = \mathcal{Z} = 100$ and Y_0 is assigned the values of S_{100} during the fitting process. At the same time, the equivalent of eq. (28) can be fitted as a base-age specific site index equation using base-age 50 years if $\mathcal{Z} = 100$, $t_0 = 50$ and Y_0 is assigned the values of S_{50} during the fitting process, i.e.,

$$Y(t, S_{50}) = \exp\left(\frac{\ln(Y_0\, 2^{\alpha+\beta\ \ln t_0-\ln Y_0(\gamma+\delta\ \ln t_0)}\, (.01t)^{\alpha+\beta\ \ln t})}{1 + (\gamma + \delta\ \ln t)\ln(.01t)}\right)$$

Furthermore, both of these equations, either with base-age 100 or with base-age 50 years, could in the application phase be used directly—without reformulation—to calculate site indexes at either of the two base-ages or any other base-age. For example, the model could be base-age 50 specific ($\mathcal{Z} = 50$) but be used with base-age 100 site indexes ($t_0 = 100$) for height predictions:

$$Y(t) = \exp\left(\frac{\ln(Y_0\, 2^{\ln Y_0(\gamma+\delta\ \ln t_0)-\beta\ \ln t_0-\alpha}\, (.02t)^{\alpha+\beta\ \ln t})}{1 + (\gamma + \delta\ \ln t)\ln(.02t)}\right)$$

Similarly, the equivalent of eq. (30) can be fitted using eq. (34) as a base-age specific site index model using base-age 100 years if $\mathcal{Z} = 50$, $t_0 = 100$ and Y_0 is assigned the values of S_{100} during a fitting process. The equivalent of eq. (30) can be fitted as a base-age specific site index model using base-age 50 years if $\mathcal{Z} = 50$, $t_0 = 50$ and Y_0 is assigned the values of S_{50} during the fitting process. Furthermore, both of these models, based on either base-age 100 or base-age 50 years, could be used directly without reformulation to calculate site indexes at either of the two base-ages or any other base-age from any height (Y_0) and age (t_0) measurements. For example, the following equation is equivalent to eq. (29) ($t = 100$) but is base-age 50 specific ($\mathcal{Z} = 50$):

$$S_{100}(t_0, Y_0) = \exp\left(\frac{\ln(Y_0\ (.02t_0)^{\ln Y_0(\gamma+\delta\ \ln t_0)-\beta\ \ln t_0-\alpha}\ 2^{\alpha+\beta\ \ln t})}{1 + (\gamma + \delta\ \ln t)\ln 2}\right)$$

Equation (34) can be used for analysis of many other fitting techniques, such as those described in Borders *et al.* (1988) and Furnival *et al.* (1990), i.e., all possible combinations of data measurements, non-overlapping growth intervals, etc. The differences in curves from different methodologies as discovered by others support Curtis' conclusion that all base-age specific methodologies as tested on his data produce different curves. This holds also for the methods described by Borders *et al.* (1988) and Furnival *et al.* (1990) and some other methodologies tested on the same data[5].

Finally, since equations (28) and (30) are, in fact, two different equations having different properties and varying by the arbitrary constants, 100 vs. 50, one may well ask the question:

[5]Personal communication: First author's correspondence with Dr. R.O. Curtis and Dr. B.E. Borders, July 13, 1990.

> What value of the constant, 100, 50 or some other, results in the best curves given an arbitrary base-age specific fitting using, say, the base-age 75 years or some other base-age?

Such a question cannot be answered with either eq. (28) or eq. (30). There is no a direct way to fit eq. (28) or eq. (30) with base-age 75 years, or other base-ages. Nor are these equations conditioned to predict heights equal to site indexes at base-age 75 years, or other base-ages.

The derivation of eq. (34) allows one to answer these and other similar questions. Equations (28) and (30) can be fitted and compared directly with each other using base-age 75 years ($t_0 = 75$; $Y_0 \equiv S_{75}$; and $\mathcal{Z} = 100$ vs. $\mathcal{Z} = 50$) or any other base-age. Moreover, an infinite number of equations similar to eq. (28) and eq. (30) with various (different than 100 or 50) constants can be analyzed simultaneously in one regression run using eq. (34) by simply defining $\mathcal{Z}$ as an estimable regression parameter. Such a parameter ($\mathcal{Z}$) can be estimated by any base-age specific regression regardless of the value of the regression base-age (t_0). Furthermore, the predicted values from any model based on this generalized equation (34) will always give height equal to site index at any base-age t_0 and for any value of $\mathcal{Z}$.

Discussion

The focus of this manuscript is on a methodology for algebraic derivation of dynamic equations that are suitable for modeling pooled cross-sectional and longitudinal data and that more flexible than other methods given in the forestry literature on dynamic equations. The equations derived can be fitted to data with any technique suitable for dynamic or fixed-base-age equations. Furthermore, these equations can be used, if desired, in ways consistent with the other more traditional fixed-base-age equations. We recommend the methodology as a tool and not as an ideology. It does not pre-empt any statistical assumptions on error structures or criteria of fitting. We do not claim that all site models must be based on dynamic equations. Yet, we have provided evidence (e.g., compare eq. (22) vs. eq. (20) and (21)) that given certain curve shapes, the dynamic equations are superior to the fixed-base-age equations and other explicit equations. They are generally more parsimonious and flexible. They will predict appropriate heights when age equals base-age and will be easier to fit with scant data or data from young trees.

We include examples of dynamic equations used for fitting base-age-independent and base-age-specific parameters. The purpose of these examples is to demonstrate the advantages of the Generalized Algebraic Difference Approach to equation derivation over more traditional approaches. These advantages arise from greater flexibility in the model analysis and its applications.

Acknowledgements We dedicate this work to Dr. J.L. Clutter who died much too early. Nonetheless, he inspired and advocated the use of implicitly defined dynamic equations in modeling of all site dependent forest characteristics.

We are grateful to Drs. I.E. Bella, J. Beck, H.E. Burkhart, R.O. Curtis, V. Lieffers, S. Northway, R.A. Monserud, S. Titus, C. Tomas, and L.V. Pienaar, for their reviews of other unpublished manuscripts describing the generalized algebraic difference methodology and to Dr. D. Tait for his feedback relating to the concept of unobservable variables

We extend special thanks to Dr. R.O. Curtis. Bob reviewed one early manuscript several times, entertained extensive discussions on related issues, and shared the noble fir data and unpublished works. With his permission, some outcomes from this collaboration are included here. It was Bob's idea, and a very helpful one, to use just one specific base model (Schumacher 1939) to illustrate the various derivations.

All help has been greatly appreciated but any shortcomings are our own doing.

Literature Cited

Amaro, A., D. Reed, M. Tome, and I. Themido. 1998. Modeling dominant height growth: Eucalyptus plantations in Portugal. For. Sci. 44: 37-46

Bailey, R.L., and J.L. Clutter. 1974. Base-age invariant polymorphic site curves. For. Sci. 20: 155–159.

Begin J., and J.P. Schutz. 1994. Estimation of total yield of douglas-fir by means of incomplete growth series. Ann. Sci. Forest. 51: 345-355

Borders, B.E., R.L. Bailey, and K.D. Ware. 1984. Slash Pine site-index from a polymorphic model by joining (splining) nonpolynomial segments with an algebraic difference method. For. Sci. 30: 411–423.

Borders, B.E., R.L. Bailey, and M.L. Clutter. 1988. Forest growth models: parameter estimation using real growth series. P. 660-667. In A.R. Ek, S.R. Shifley, and T.E. Burk (eds.). Forest growth modelling and predicion. Vol. 2. Proc. IUFRO conf. USDA For. Ser. Gen. Tech. Rep. NC-120. 1149 p.

Cao, Q.V. 1993. Estimating coefficients of base-age-invariant site index equations. Can. J. For. Res. 23: 2343-2347.

Cao, Q.V., V.C. Baldwin, Jr., and R.E. Lohrey. 1997. Site index curves for direct-seeded loblolly and longleaf Pines in Louisiana. NJAF 21: 134-138.

Cieszewski, C.J. 1994. Development of a variable density height-growth-model through defining multidimensional height growth spaces. Ph.D. Thesis, U. of A., Edmonton, AB, CA. 72p

Cieszewski, C.J., and I.E. Bella. 1989. Polymorphic height and site-index curves for lodgepole pine in alberta. Can. J. For. Res. **19:** 1151-1160.

Cieszewski, C.J., and I.E. Bella. 1993. Modeling density-related lodgepole pine height growth, using Czarnowski's stand dynamics theory. Can. J. For. Res. **23:** 2499–2506.

Clapham C. 1996. The concise oxford dictionary of mathematics. Second Edition. Oxford New York. Oxford University Press. p.313.

Clutter, J.L., J.C. Fortson, L.V. Pienaar, G.H. Brister, and R.L. Bailey. 1983. Timber management. John Wiley and Sons, Inc. New York. 333p.

Clutter, J.L., W.R. Harms, G.H. Brister, and J.W. Rheney. 1984. Stand structure and yields of site-prepared loblolly Pine plantations in the Lower Coastal Plain of the Carolinas, Georgia, and north Florida. USFS Gen. Tech. Rep. SE-27. 173 pp.

DuPlat, P. and M. Tran-Ha. 1986. Modeles de croissance en hauteur dominante pour le h être, le sapin pectiné et le pin sylvestre (dans le Massif de l'Aigoual). Document n^o 86.1. Office National de Forêts. France.

Furnival, G.M., T.G. Gregoire, and H.T. Valentine. 1990. An analysis of three methods for fitting site-index curves. For. Sci. **36:** 464–469.

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

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

Gellert, W., H. Küstner, M., Hellwich, and H. Kästner. 1977. The VNR concise encyclopedia of mathematics. Van Nostrand Reinhold Company. 760 p.

Hosfeld, J.W. 1822. Mathematic für forstmänner, ökonomen und cameralisten. Gotha 4. Bd., S. 310. (In German).

Lappi, J., and R.L. Bailey. 1988. A height prediction model with random stand and tree parameters: an alternative to traditional site-index methods. For. Sci. 34: 907–927.

Lenhart J.D. 1968. Yield of old-field loblolly pine plantations in the Georgia Piedmont. Ph.D. Thesis. School of Forest Resources, The University of Georgia, Athens, Georgia (unpublished).

Lenhart J.D. 1972. An alternative procedure for improving height/age data from stem analysis. For. Sci. 18:332.

McDill, M.E., and R.L., Amateis. 1992. Measuring forest site quality using the parameters of a dimensionally compatible height growth function. For. Sci. 38: 409-429.

Osborne, J.G. and F.X. Schumacher. 1935. The construction of normal-yield and stand tables for even-aged timber stands. J. Agric. Res. 51:547—564.

Ramirez, M.H., R.L. Bailey, and B.E. Borders. 1987. Some implications of the algebraic difference approach for developing growth models. p. 731-738, In Proc. “IUFRO Forest Growth Modelling Symposium”, Minneapolis, Minn. August 24–28, 1987.

Schnute, J. 1981. A versatile growth model with statistically stable parameters. Can. J. Fish. Aquat. Sci. 38: 1128–1140.

Schumacher, F.X. 1939. A new growth curve and its application to timber yield studies. J. For. 37: 819–820.

USDA. 1929. Miscellaneous publication No. 50. Washington, D.C.