ELSEVIER

Forest Ecology and Management 97 (1997) 165–172

Forest Ecology and Management

Height growth modeling using second order differential equations and the importance of initial height growth

Rolfe A. Leary [a,*], Kevin Nimerfro [a], Margaret Holdaway [a], Gary Brand [a], Tom Burk [b], Randall Kolka [b], Ann Wolf [b]

[a] *USDA Forest Service, North Central Forest Experiment Station 1992 Folwell Ave., St. Paul, MN 55108, USA*
[b] *University of Minnesota, St. Paul, MN 55108, USA*

Accepted 12 December 1996

Abstract

In a study of height growth patterns of quaking aspen (*Populus tremuloides* Michx.) in Minnesota, Wisconsin and Michigan, we represented the height–age pattern using a second order ordinary differential equation with environmentally governed parameters. Solving a second order differential equation that has been converted to a system of two first order equations requires knowledge of, or information on, initial conditions for both state variables, height and height growth. We used the natural boundary condition, $h(t=0)=0$. Initial conditions for the second state variable, height growth, were estimated when fitting the equations to observations. This paper reports our research to predict the initial height growth from stem analysis data, and to assess the sensitivity of predicted height to differences in initial condition estimates. We found initial height growth to be the single most important 'parameter' to be estimated in our model. © 1997 Elsevier Science B.V.

Keywords: Height growth; Second order differential equations; *Populus tremuloides*; Lake States

1. Introduction

Height is an important property of forest trees and stands from several points of view: species succession (taller trees may spread the most propagules and have a reproductive advantage), forest mensuration (percent of height is used as spacing indicator), and site assessment, where theory and practice have claimed that tree height is an important signal of growing environment suitability (height–age, height–diameter relations can be used as indicators of site quality).

Representing the height–age relation is more of a modeling challenge than one might expect because when empirical data are looked at closely, the pattern can be much more irregular than is described by Spurr (1952). We conjectured that tree height growth could be represented by identifying a simple basic model that makes use of more advanced mathematical methods (integro-differential equations), with focus shifting from getting flexibility out of the model to identifying the environmental factors governing model parameters.

Results for a larger modeling study that focuses on determining environmental variables governing model parameters are reported elsewhere (Leary et al., in process). We report here only a study of the importance of the first year's height growth in pre-

---

[*] Corresponding author. Tel.: 612-649-5177; Fax: 612-649-5285; E-mail: RLeary@forestry.umn.edu.

PII S0378-1127(97)00097-2

dicting heights for later ages. We describe below the overall modeling strategy primarily to give background on how the estimates of initial height growth were developed.

2. Methods

We tested the modeling approach on quaking aspen, *Populus tremuloides* Michx., growing in the Lake States of the United States (Minnesota, Wisconsin, Michigan). Most data came from quaking aspen growing in Minnesota. Quaking aspen reproduces from root suckers when mother stands are clearcut, but can also reproduce from stump sprouts and seed. The latter modes are much less common, however.

The base model we conjectured is the simple integro-differential equation

$$\dot{h} = ct - ah - b\int_{u=0}^{u=t} f(h(u))\,du \qquad (1)$$

(Hamlin and Leary, 1987). (We use the standard single dot notation to denote the first derivative with respect to time. Double dots denote the second derivative with respect to time.) Eq. (1) can be differentiated to remove the integral from the right-hand-side, and when the function $f(h(u))$ is the identity function, the result is a linear second order equation:

$$\ddot{h} = c - a\dot{h} - bh. \qquad (2)$$

Following standard practice, we re-express Eq. (2) as an equivalent simultaneous system of first order differential equations:

$$\dot{h} = g$$

$$\dot{g} = c + a\dot{h} + bh \qquad (3)$$

where: h designates total tree height; g designates tree height growth; a, b, c denote base model parameters.

Hypothesizing that the base model parameters are simple linear functions of environmental variables, we add indexing for observation times (k), trees (j), and sites (i), and associate environmental variables with particular base model parameters:

$$\dot{h}_{ijk} = g_{ijk}$$

$$\dot{g}_{ijk} = (c_0 + c_1V_1(t)) + (a_0 + a_1V_2(t))\dot{h}_{ijk} + (b_0 + b_1V_3)h_{ijk} \qquad (4)$$

where $V_1(t)$ is an environmental variable representing the annual heat status of the site, $V_2(t)$ is an environmental variable representing the annual moisture status of the site, V_3 is a site variable representing the nutrient status of the site, assumed to be constant over the life of the tree, but different among sites, a_0, a_1, b_0, b_1, c_0, c_1 denote parameters of the hypothesized linear relations between base model parameters and environmental variables, h_{ijk} designates total tree height (m) at age k, for tree j, on site i, and g_{ijk} designates height growth (m) at age k, for tree j, on site i.

We tested other functional relations between model parameters and environmental variables, and found simple linear functions adequate, given the limitations of height–age data. The unknowns are the parameters in Eq. (4), and the initial height growth, g_0. The boundary conditions are the heights of different trees on different plots at different ages, h_{ijk}. Because of the large computational load we found it necessary to convert our model from first order differential equations to first order forward difference equations.

We use a least squares fit criterion:

$$\mathrm{Min}\{SS \mid \alpha_1 \alpha_2 \alpha_3 \alpha_4\}$$

where

$$SS = \sum_{i=1}^{S} \sum_{j=1}^{T} \sum_{k=1}^{M} \left(h_{ijk}(\text{observed}) - h_{ijk}(\text{predicted})\right)^2, \qquad (5)$$

where S denotes the number of sites, T the number of trees, and M the number of observations, and where

h_{ijk}(predicted)

= solution to inhomogeneous form of [4]

+ α_1 1st solution to homogeneous form of [4]

+ ... + α_4 4th solution to homogeneous form of [4]

(For this analysis we chose to ignore correlation among trees and measurements on the same tree. Our reasons are given in Section 5)

The α's are the superposition constants that result from linearizing the model as done by Bellman and Kalaba (1965). Example forestry applications are given in Leary (1970). Parameters of Eq. (4) and their variances can be derived from the variance–covariance matrix for the α's, using standard linear least squares regression methods. Significance tests for these parameters also follow standard methods for linear regression (Compte, 1994).

Two categories of estimation problems arose: first, what are the environmental variables governing the base model parameters, and second, what factors govern the initial height growth (which must either be measured or estimated)? Our scientific hypotheses for the first problem began by focusing on base model parameters in Eq. (3). The hypotheses were: a is governed by moisture availability; b is governed by nutrient availability; c is governed by heat.

Resolving the second problem is the primary purpose of this paper. As noted above we devote some space to the first problem to give context to our proposed solution, and to show how estimates of initial height growth were developed.

3. Materials

To test the above hypotheses we assembled existing stem analysis data for quaking aspen from a study by Schlaegel (1975), and a more recent study reported on by Alban et al. (1991). Field and office procedures are reported in these publications. An additional 13 plots were established across northern Minnesota to span a range of climate and soil fertility conditions (Miller and Cooper, 1993). On these additional plots four trees growing within 50′ of the plot center were felled and stem disks taken at 2-ft intervals on each tree. Tree discs were dried, sanded with three grades of sandpaper, wetted with water, scanned using the MacDendro scanning package (Guay et al., 1992), and ring locations manually identified from a computer monitor. The climate record was not made available to the persons scanning discs to avoid the bias that would result, for example, from assigning years to rings based on apparent precipitation events. We applied Carmean's method for estimating terminal bud location between discs to all data sets (Dyer and Bailey, 1987). Sixty seven trees were available for analysis from 22 locations. About 20 height–age pairs were available for each tree, but this depended on height interval between disc sampling. All height–age data sets were truncated to the 0–50 age interval.

Soil physical and chemical data were assembled for all data sets. Unfortunately, the properties sampled were not uniform across sets. We estimated missing physical and chemical properties from those properties that were available for each study using relationships in the published literature. We spatially interpolated the historical climate records at the plot locations from neighboring meteorological station records using kriging methods (Holdaway, 1994). Climate properties interpolated were mean annual temperature, annual precipitation, and annual growing degree days. We adapted a method (first developed by Black (1966) and modified by Grigal and Bloom (1985)) to estimate monthly actual evapotranspiration at the plot location.

To test the utility of the nutrient synecological coordinates to proxy for the nutrient status of the site, we collected plant lists for the plot locations (Bakuzis, 1959, Bakuzis and Kurmis, 1978). Locations of our 13 plots were known exactly, however, only approximate synecological coordinates could be computed in the Schlaegel (1975) and Alban et al. (1991) studies because the location of the growing site either was not known exactly or the stand had been clearcut and converted to another species.

We projected the height growth of individual trees on each plot by solving the associated difference equation using the historical record to calculate the values of each environmental variable for each year and, in term, calculating a different value of base model parameters a and c for each year (see below).

4. Results

When we fit the model to the 67 trees individually, mean deviations were quite small (Fig. 1a). However, we had used 268 base model parameters (67 estimates for a, b, c, and g_0) and no measured environmental variables. When we used a single

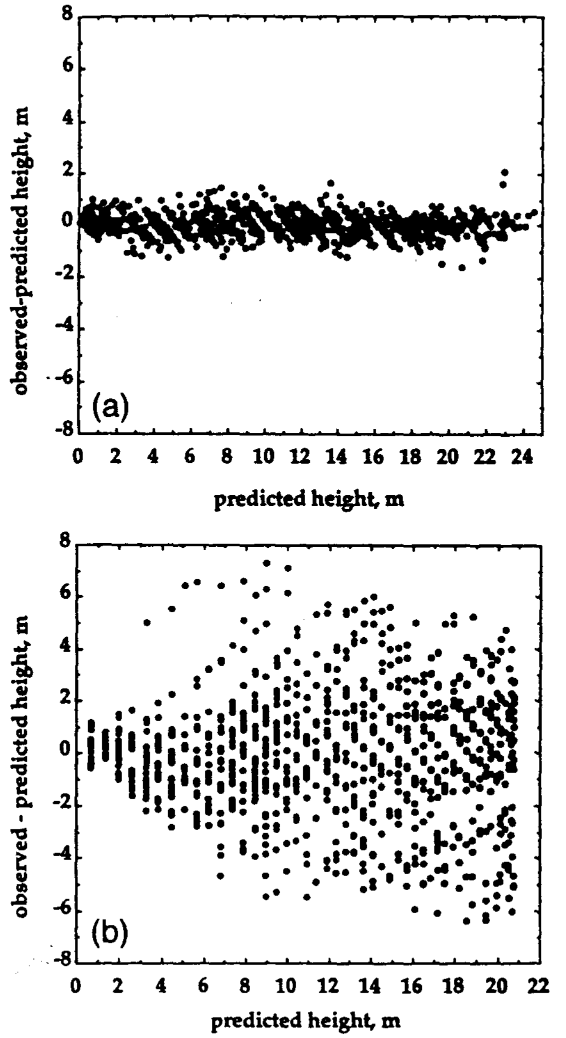

Fig. 1. (a) Residual pattern when Eq. (3) was fit to data for each tree individually. Number of parameters estimated per tree was four, for a total of 268. (b) Residual pattern (observed - predicted) when Eq. (3) was fit to pooled data for all trees. Number of parameters estimated was four.

equation to represent an average height growth pattern, the deviations were, of course, large. However, we had only used four parameters (Fig. 1b).

To approach the narrow residual pattern of Fig. 1a we included information on the annual climate and soil conditions at the growing site and associated these conditions with base model parameters *a*, *b*, *c* (cf. Eq. (4) and Eq. (2)). Analysis of the data supports our hypothesis that base model parameter *c* is governed by heat at the growing site expressed as growing degree days above 5°C, and that base model parameter *a* is governed by estimated actual annual evapotranspiration. We also determined that while base model parameter *b* has a statistically significant relation with the nutrient synecological coordinate, it is not practically significant, once growing degree days and estimated actual evapotranspiration are associated with base model parameters *c* and *a*, respectively. Our nutrient data were probably inadequate given the very high spatial variability of soil chemical properties in our region (Alban, 1974, Mroz and Reed, 1991).

Significantly, we found that in predicting future heights the initial height growth was the most important single 'parameter' in the model. Our discovery was gratifying on one hand because initial height growth can be measured with simple equipment. However, it also caused concern because it is known that initial tree height growth can be extremely variable (Palik and Pregitzer, 1995).

We turned our attention to finding simple ways to estimate the important parameter, $\hat{g}_0$. Of course, we were not present to actually measure the early height growth of each tree in our data set, so we don't have actual measured tree heights to use as a dependent variable. Our 'observations' of early heights were estimates based on Carmean's method of interpolating the location of the terminal bud between discs cut at either 2- or 4-ft intervals above a one half foot stump. The interpolation process made these 'observations' very lumpy, so we abandoned using the interpolated initial height growth estimates, using instead the 67 estimates of the initial condition, $\hat{g}_0$, as the dependent variable.

We tested two approaches to estimating: $\hat{g}_0$: (a) base estimates on the environment in the year of regeneration and (b) base estimates on periodic height growth from trees used in our sample.

4.1. Initial height growth and environmental conditions during the regeneration year

It would be ideal if we could predict initial height growth from environmental variables, either by site conditions or climate conditions during the regeneration year. We had only basic information on climate during the year of regeneration from the interpolated historical record. Of course, we had no way of knowing the soil conditions at the time of regeneration, although we assumed that physical properties were essentially the same as when the trees were felled for analysis. Graphical analyses showed that estimated initial height growth was not well correlated with any environmental variable with the possible exception of a height growth peak at about 14.1°C mean annual temperature.

We then tested two sources of periodic height growth information to predict early height growth. Sources were mean annual height growth over the first few years of the tree's life, and height at 50 yrs.

4.2. Early growth estimated from periodic height growth from early growth measurements

We tested a strategy based on taking height measurements in young stands spanning a few years, computing a mean height growth per year, and using the computed mean height growth to predict initial height growth. With the predicted initial height growth, one can then solve the height growth equations (Eq. (4)) to predict height at all ages. As long as the wait is not too long before the just described sequence of steps can be reliably carried out, it may be a satisfactory way of estimating heights in older stands.

The primary question we examined was: between what ages or heights is the mean height growth best correlated with initial height growth estimates, $\hat{g}_0$? Because we were working with quaking aspen, we could not do a whorl-based analysis as applied by Alban (1972, 1979) when he developed site intercept methods for red pine. However, the principles we applied were similar.

We analyzed disks cut from the stem at different intervals among our data sources (2 ft, 4 ft, 1 m) to determine age and to estimate the location of the terminal bud between disks. We calculated mean height growth using ground level and breast height as reference points, and then calculated mean annual height growth between disks 1 and 3, 2 and 4, 3 and 5 above ground line, as well as mean annual height growth based on the first and third, first and fourth, and first and fifth disks above breast height. The six mean height growth estimates were calculated for each individual tree, converted to plot average values and graphed against the mean of model-based estimates of initial height growth, $\hat{g}_0$ for each plot. Ground line as the reference height gave a smaller variance than when breast height was used as the reference height. For our observations, the relation between estimated height growth based on ages and heights at the 2nd and 4th disks cut above ground line gave estimates of g_0 with the lowest variance (Fig. 2). There were on average, about 5-yr difference between the age at the second and fourth disks, and about 3-m difference in height. Application of the just described procedure presumes one can, in sucker stands over 5 yrs of age, visually approximate 5 yrs of past height growth. Mean annual height growth is used in the following equation to predict the estimated initial height growth:

$$\hat{g}_0 = 1.394 \left(1 - e^{-0.903\,(\text{mean annual height growth between disk 2 and 4})}\right) \tag{6}$$

Adjusted R squared = 0.712, Standard error of estimated growth = 0.10 m (Wilkinson, 1990).

The sample point to the right in Fig. 2 is for trees in our sample from Pike Bay West (Minnesota). Their very rapid growth lead us to suspect that they

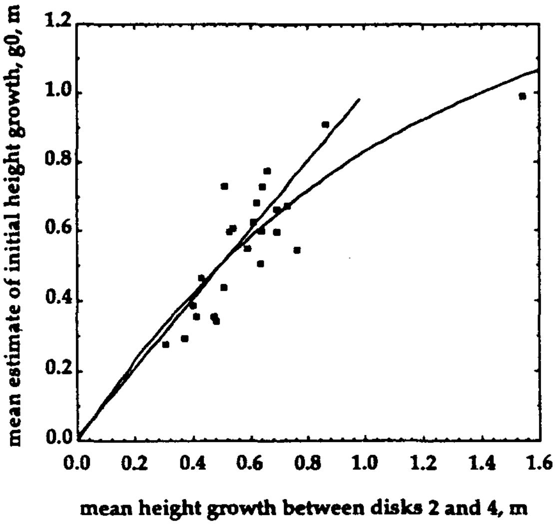

Fig. 2. Relation between early growth rate between the 2nd and 4th disks cut from trees, and estimates of initial height growth, $\hat{g}_0$. Eq. (3).

might have been of stump sprout origin, rather than root sucker. If the single point corresponding to Pike Bay West is removed from the dataset, the relation between annual height growth between the 2nd and 4th disks, and $\hat{g}_0$, is linear (Eq. (7)):

$$\hat{g}_0 = 0.0107 + 0.990 \text{ (mean annual height growth between disks 2 and 4)}, \quad (7)$$

Adjusted R squared = 0.648, Standard error of estimated growth = 0.099 m (Velleman, 1992).

Because the slope is not significantly different from 1.0, and the intercept term is nonsignificant, the initial condition for g_0 can be directly measured in the field by visiting, on average, 6-yr-old stands, measuring the height growth in the 5 prior years, and dividing by 5.

4.3. Early growth estimated from 50-yr height growth measurements

The second method for estimating initial height growth simulates use of mother stand height at an index age (e.g., 50 yrs) to predict initial height growth in the daughter stand. The method presumes that stem analysis has been carried out on a sample of felled mother stand trees if felling occurred at an age different than the index age. For the test here, the subject trees are also the trees for which we are estimating initial height growth. There is obvious circularity in the second method as we apply it, so our findings are probably unrealistically accurate. It should be imagined to be applied to mother stands for prediction of initial height growth in daughter stands.

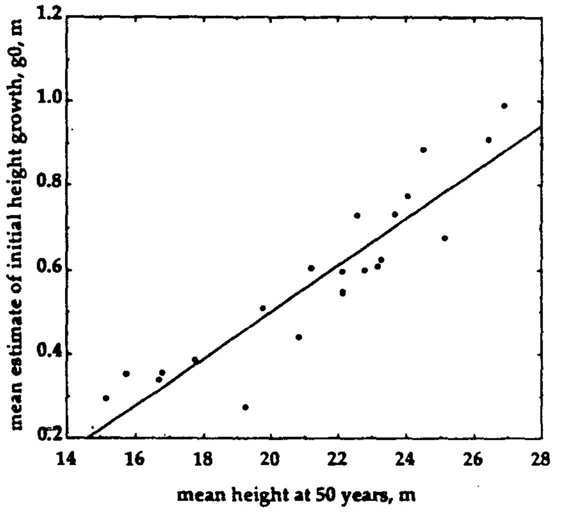

Fig. 3. Relation between plot mean initial height growth estimate and observed mean tree height at 50 yrs.

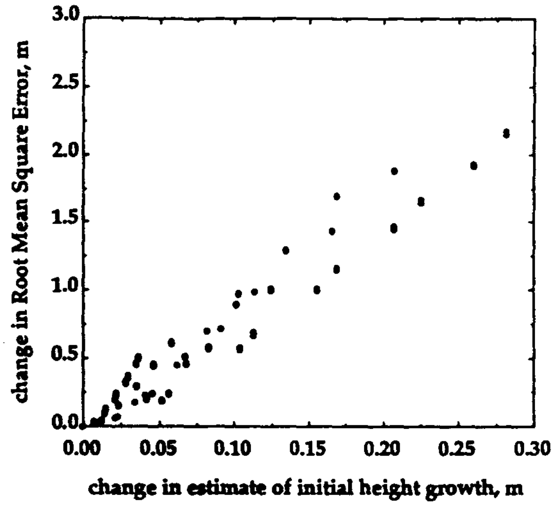

Fig. 4. Relation between systematically perturbed estimates of initial height growth, g_0, and the root mean square error of observed from predicted heights. Note: a 0.2 m perturbation in estimated initial height growth resulted in a root mean square error increase of about 1.45 m.

We used a simple linear model to approximate the relation between the estimated initial height growth, g_0, and height at the index age. Recall, our model was based on estimates of g_0 when parameter b was based on the site nutrient synecological coordinates. Plot means were used for both g_0 and tree height (Fig. 3). The final model is:

$$\hat{g}_0 = -0.6030 + 0.0552 \text{ (mean height at 50 yrs)} \quad (8)$$

Adjusted R squared = 0.83, standard error of estimated growth = 0.083 m (Velleman, 1992).

4.4. Sensitivity of model performance to errors in estimating initial height growth

The sensitivity of the model to errors in estimating initial height growth was tested on nine trees selected to have a wide range in estimated initial height growth. Sensitivity was judged by systematically changing the initial height growth in increments of 5% both above and below the estimated g_0. Each tree's growth was projected, starting from the altered initial growth, using the actual annual historical record of growing degree days, the estimated annual record of actual evapotranspiration, the nutrient synecological coordinate, and the final estimates of parameters a_0, a_1, b_0, b_1, c_0, and c_1 for each specific tree. Three measures of deviation of predicted from observed heights were calculated: mean deviation, mean absolute deviation, and root mean squared error (RMSE). Relations among initial growth change and the deterioration of mean absolute error are shown in Fig. 4. The regression intercept value is not statistically significant. The slope was 7.25 m per meter of change in initial height growth.

Our findings from the sensitivity analysis confirm our previous finding that initial height growth is extremely important in predicting heights through the life of the tree.

5. Discussion

On the basis of our study results we conclude that having a good estimate of initial quaking aspen height growth is more important for estimating future heights than having a complete history of either heat sums or actual evapotranspiration spatially interpolated to the plot location. However, getting estimates of initial height growth in quaking aspen, after the fact, is not easy.

We found that early height growth could not be predicted well using environmental information available to us some 60–90 yrs following the regeneration event. So, although we cannot claim conclusively that initial height growth cannot be predicted from climate and site properties, we feel justified in our claim that evidence available in this study does not support reliable prediction of initial height growth from climate and site data in the regeneration year.

When we looked at a short and long period of height growth to use to predict, actually to retrodict, initial height growth, predictability depends on the length of the period. Using a very long period was superior to using a shorter period. Thus, gathering total heights in mother stands at harvest time may be a good source of information to predict height growth at regeneration. We can only suppose that mother stand heights would be useful because in our study we used the heights at age 50 yrs of the actual felled trees to predict their initial growth 50 yrs previously.

From a practical point of view, prediction of initial height growth from periodic early height growth (say over 5 or 6 yrs) would be preferable. Our analysis showed the approach, similar to the site intercept concept, is possible, although the results are not as precise as when 50-yr heights are used. In general, getting good estimates of initial height growth may be difficult, especially if quaking aspen is over about 6 yrs of age. (One can estimate previous year's height growth from morphological features of the stem for a few years down from the leading shoot, but probably not over about 5 yrs.) Splitting the log longitudinally is one alternative, but we could have used this approach only on trees from our study. The Schlaegel (1975) and Alban et al. (1991) study trees had been left in the woods to rot about 20 yrs previously. Further, it is not clear whether the terminal bud location would be evident in a split quaking aspen lower bole. Periodic, temporary inventories of regenerating aspen stands would seem to be the most reliable source of the needed information.

Our findings on the sensitivity of height prediction errors through the life of the tree to errors in estimating initial height growth were disturbing, to say the least. Very small errors in estimating initial height growth can be multiplied as much as seven fold in predicting later tree heights.

Recent growth model research has focused on problems in parameter variance estimation in the presence of serially correlated errors. In this study, reduction of serially correlated errors was possible by using the environmental variables of heat sums and estimated actual evapotranspiration as forcing functions on model parameters. In some cases the reduction in serial correlation was dramatic, while in others, there was less reduction. In retrospect we think our negative results were related to errors in ring counts, known to be difficult in quaking aspen.

A common approach to the problem we faced is to use a two-stage process, wherein a model is first fit to tree or plot observations individually, and then the sets of estimated parameters, say sets of our base model parameters a, b, c, are related to environmental variables (e.g., Ferguson and Leech, 1978). We did not require two stages to reach approximately the same end point. Our fitting approach allowed us to fit the model in Eq. (4) to the 67 trees simultaneously, each tree being 'driven' by its own unique historical record of climate and site conditions, to estimate values of a_0, a_1, c_0, c_1, b_0, and b_1 that are common to all trees, at the same time we estimated 67 g_0 values specific to individual trees.

While it is commonly accepted that future growth is influenced greatly by current size - this is the basis for nearly all autonomous ordinary differential equations of growth, e.g., von Bertalanffy and many others - our finding is that for quaking aspen, future growth is greatly dependent on first year size, not just current size. As B. Elfving states in review comments, 'good initial height growth... can have a lasting effect on the productivity.'

In terms of model selection, our finding suggests that for quaking aspen, and perhaps other species of sprout origin, the appropriate height growth model is not the integro-differential equation we used, but a delay-differential equation, with a variable delay so that the first year height is always the delay term, rather than the usual formulation of a delay-differential equation with a fixed delay. Such a delay-differential equation would have the form $dh/dt = f(h, h(1))$, rather than the usual form $dh/dt = f(h, h(t-k))$, where k is the fixed delay.

References

Alban, D., 1972. An improved growth intercept method for estimating site index of red pine. US For. Serv. North Cent. For. Exp. Stn. Res. Pap., NC-80. 7 p.

Alban, D., 1974. Soil variation and sampling intensity under red pine and aspen in Minnesota. US For. Serv. North Cent. For. Exp. Stn. Res. Pap., NC-106. 10 p.

Alban, D., 1979. Estimating site potential from the early growth of red pine in the Lake States. US For. Serv. North Cent. For. Exp. Stn. Res. Pap., NC-166. 7 p.

Alban, D., Perala, D., Jurgensen, M., Ostry, M., Probst, J., 1991. Aspen ecosystem properties in the Upper Great Lakes. US For. Serv. North Cent. For. Exp. Stn. Res. Pap., NC-300. 47 p.

Bakuzis, E., 1959. Synecological coordinates in forest classification and in reproduction studies. Ph.D. Dissertation. University of Minnesota. 244 p.

Bakuzis, E., Kurmis, V., 1978. Provisional list of synecological coordinates and selected ecographs of forest and other plant species in Minnesota. Staff Series Paper Number 5. Department of Forest Resources, University of Minnesota. 31 p.

Bellman, R., Kalaba, R., 1965. Quasilinearization and nonlinear boundary-value problems. Elsevier, New York. 206 p.

Black, P., 1966. Thornthwaite's Mean Annual Water Balance. Tech. Pub. No. 92 of the State University College of Forestry, Syracuse, New York.

Compte, W., 1994. A Monte Carlo comparison of two approaches for calibrating differential equation models. A Plan B paper submitted to the Faculty of the Graduate School of the University of Minnesota. 143 p.

Dyer, M., Bailey, R., 1987. A test of six methods for estimating tree heights from stem analysis data. Forest Science 33, 3–13.

Ferguson, I.S., Leech, J.W., 1978. Generalized least squares estimation of yield functions. Forest Science 24, 27–42.

Grigal, D., Bloom, P., 1985. Modification of the model of soil acidification. Report to Northern Stares Power Company, Minneapolis, MN.

Guay, R., Gagnon, R., Morin, H., 1992. A new automatic and interactive tree ring measurement system based on a line scan camera. For. Chron. 68, 138–141.

Hamlin, D., Leary, R., 1987. Methods for using an integro-differential equation as a model of tree height growth. Can. J. For. Res. 17, 353–356.

Holdaway, M., 1994. The microcomputer scientific software series 9: User's Guide to Geo-CLM: Geostatistical Interpolation of the Historical climatic record in the Lake States. US For. Serv. North Cent. For. Exp. Stn. Gen. Tech. Rep., NC-167. 34 p.

Leary, R., 1970. System identification principles in studies of forest dynamics. US For. Serv. North Cent. For. Exp. Stn. Res. Pap., NC-45. 38 p.

Miller, M., Cooper, T., 1993. Legislative Committee on Minnesota Resources Project Report: Impacts of intensified forest management and atmospheric change on nutrient cycling and tree species suitability. Objective B: Estimating state wide impacts. Staff Report, Department of Soil, Water, and Climate, Univ. of Minnesota, St. Paul.

Mroz, G., Reed, D., 1991. Forest sampling efficiency: Matching laboratory analyses and field sampling procedures. Soil Sci. Soc. Am. J. 55, 1413–1416.

Palik, B., Pregitzer, K., 1995. Variability in early height growth rate of forest trees: implications for retrospective studies of stand dynamics. Can. J. For. Res. 25, 767–776.

Schlaegel, B., 1975. Estimating aspen volume and weight for individual trees, diameter classes, or entire stands. US For. Serv. North Cent. For. Exp. Stn. Gen. Tech. Rep., NC-20. 16 p.

Spurr, S., 1952. Forest inventory. Ronald Press, New York. 476 p.

Velleman, P., 1992. Data Desk Statistics Guide, Volume II. Data Description, Inc. Ithaca, New York.

Wilkinson, L., 1990. SYGRAPH: The System for Graphics. Evanston, Illinois. SYSTAT, Inc. 547 p.