Technical Report 91.15

COMPARISON OF NONLINEAR HEIGHT-DIAMETER FUNCTIONS FOR MAJOR ALBERTA SPECIES

by

Shongming Huang and Stephen J. Titus
and
Douglas P. Wiens

Comparison of nonlinear height-diameter functions for major Alberta species

Shongming Huang and Stephen J. Titus[1]

Department of Forest Science

University of Alberta, Edmonton, Alberta, Canada T6G 2H1

and

Douglas P. Wiens

Department of Statistics and Applied Probability

University of Alberta, Edmonton, Alberta, Canada T6G 2G1

Phone: 403-492-1435 or 403-492-2899 or 403-492-4406

[1]Author to whom all correspondence should be addressed.

Abstract. Based on a data set consisting of 13,489 felled trees for 16 different species, 20 nonlinear height-diameter functions were evaluated for major Alberta species. Because of the problem of unequal variations for the dependent variable height, all functions were fitted using weighted nonlinear least squares regression ($w_i = 1/DBH_i$). The examination and comparison of the weighted mean squared errors, the asymptotic t-statistics of the parameters, and the plots of studentized residuals against the predicted height show that many concave and sigmoidal functions can be used to describe the height-diameter relationships. The sigmoidal functions such as the Weibull-type function, the modified logistic function, the Chapman-Richards function, and the Schnute function generally gave most satisfactory results.

Introduction

Predicting total tree height based on observed diameter at breast height outside bark is routinely required in practical management and silvicultural research work (Meyer, 1940). The estimation of tree volume, as well as the description of stands and their development over time rely heavily on the accurate height-diameter functions (Curtis, 1967). Many growth and yield models also require height and diameter as two basic input variables, with all or part of the tree heights predicted from measured diameters (Burkhart et al., 1972; Curtis et al., 1981; Wykoff et al., 1982). In the cases where the actual measurements of height growth are not available, height-diameter functions can also be used to indirectly predict height growth (Larsen and Hann, 1987).

Curtis (1967) summarized a large number of available height-diameter functions and using the Furnival's index of fit (Furnival, 1961), compared the performance of the linear functions based on the second-growth Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco) data. Since then, many new height-diameter functions have been developed. With the relative ease of fitting nonlinear functions and the nonlinear nature of the height-diameter relationships, nonlinear height-diameter functions have now been widely used in height predictions (Kozak and Yang, 1978; Schreuder et al., 1979; Curtis et al., 1981; Wykoff et al., 1982; Wang and Hann, 1988; Farr et al., 1989).

For major Alberta species, this study compared most of the published nonlinear height-diameter functions as well as some functions not previously applied to the height-diameter relationship. Although the main purpose of the study was to fit the most appropriate height-diameter functions for major Alberta species, the principles and procedures used here are generally applicable for other nonlinear model comparisons.

The data

Alberta Forest Service (AFS) provided 13,489 felled-tree data for this analysis. The data were collected over the last two decades and the sampling trees were randomly selected throughout the inventory areas of the province to provide representative information for a variety of densities, heights, species compositions, stand structures, ages, and site conditions. The data set was initially used for developing individual tree volume equations. It contains records of many different variables for individual trees and their surrounding environment. A detailed description of how the data are collected and recorded can be found in Alberta Phase 3 Forest Inventory: Tree Sectioning Manual (AFS, 1988). Two variables available from the records, diameter at breast height (DBH) outside bark and total tree height for each tree, were selected to be used in this analysis.

The 13,489 trees include 16 different species. To facilitate the analysis, species are classified into different species groups according to their similarity, importance, management objectives, and number of observations (Table 1). Summary statistics including the mean (Mean), minimum (Min), maximum (Max), and standard deviation (Std) for total tree height and DBH by species group are shown in Table 2.

Functions selected for comparison

The selection of the height-diameter functions was based on the examination of the height-diameter relationship as revealed by plotting total tree height against DBH for various species groups. Two typical examples for white spruce (*Picea glauca* (Moench) Voss) and aspen (*Populus tremuloides* Michx.) are shown in Figure 1 and Figure 2. It is clear that the height-diameter relationship for white spruce (Firure 1) has a typical sigmoidal shape, with an inflection point occuring in the lower portion of the data points. On the other hand, the shape of the height-diameter relationship for aspen (Firure 2) may be regarded as either concave or sigmoidal, with no apparent inflection point. Hence, both the typical concave functions and the sigmoidal functions were selected for evaluations. Additional nonlinear functions that are common in biological studies were also selected by considering the plots of height versus DBH compared to the typical graphs of the various functions. Table 3 provides a complete list of the selected functions. Notice some of the functions (such as 1 and 6) often appear in transformed forms, and the dependent variable may take the form of H-1.3 (Curtis, 1967). The quadratic height-diameter functions, first presented by Trorey (1932) and advocated by Ker and Smith (1955) and previously used in the Pacific Northwest (Staebler, 1954) and British Columbia (Watts, 1983), were not considered because extrapolation of the functions often leads to unrealistic height predictions.

Methods

All 20 height-diameter functions have only one independent variable so the general form of the functions can be written as:

$$[1] \quad H_i = f(D_i, \theta) + \varepsilon_i$$

where D_i is the ith observation of the independent variable DBH, θ is a vector $[\theta_1, \theta_2, ..., \theta_p]$ of p parameters, $\theta_1, \theta_2, ..., \theta_p$ correspond to a, b, c, d in the selected functions and the maximum $p = 4$; H_i is the ith observation of the dependent variable height; ε_i is the random error term.

A fundamental nonlinear least squares assumption is that the error terms ε_i are independent and identically normally distributed with zero mean and constant variance σ^2, that is, $\varepsilon_i \sim N(0, \sigma^2)$. However, in many forestry situations there is a common pattern of increasing variation for larger values of the dependent variable. This is clearly evident from the scatter plots of height versus DBH in Figure 1 and Figure 2, where the values of the error ε_i are more likely to be small for small DBH and large for large DBH. When the problem of unequal error variances occurs, weighted nonlinear least squares (WLS) is applied with the weights selected to give less importance to observations with large errors. The estimated θ for WLS is obtained by minimizing:

$$[2] \quad SSE(\theta) = \sum_{i=1}^{n} w_i (H_i - \hat{H}_i)^2$$

where H_i is the observed total tree height and $\hat{H}_i$ is the predicted total tree height by the fitted function; w_i is the weight corresponding to the point (D_i, H_i) and theoretically, is taken to be inversely proportional to the variance of the error term.

The WLS estimates of the parameters use an iterative process with a starting value of θ chosen and continually improved until the weighted error sum of squares SSE(θ) is minimized. Although there are many different methodologies that are available for obtaining the nonlinear least squares estimates (Gallant, 1987; Seber and Wild, 1989), in general, one should be aware that the estimates of the parameters are *not* unbiased, normally distributed, or minimum variance; rather, they achieve these properties only asymptotically (Ratkowsky, 1983, 1990; Gallant, 1987; Rawlings, 1988). As shown in Table 2, sample size for all species groups is reasonably large and therefore asymptotic test statistics are applicable.

It should be noted that the use of the weighted nonlinear least squares changes the estimates of the parameters and the standard errors of the estimates relative to the values obtained in the absence of weighting (Ratkowsky, 1990). The interpretations of the weighted statistics are not as straightforward as those in the cases of without weighting (Freund and Little, 1986; Carroll and Ruppert, 1988). However, comparison of the fit statistics for various functions can be made if the same weight is consistently used in all the function fittings and the same nonlinear least squares iteration procedure such as the Gauss-Newton method is used.

The use of the WLS requires a known weight w_i. In many practical applications, however, this weight may not be readily available so an estimate based on the results of an unweighted least squares fit is often necessary. Although there are many different procedures that are available for approximately estimating the weight or implementing the generalized nonlinear least squares techniques (Gallant, 1987; Judge et al., 1988), a simpler procedure that is based on the analysis of the studentized residuals can be equally efficient.

Studentized residuals are the scaled version of residuals that are obtained by dividing each residual by its standard error:

$$[3]\quad r_i^* = \frac{r_i}{\sqrt{MSE(1-h_{ii})}}$$

where r_i^* is the studentized residual, r_i is the residual, MSE is the mean squared error defined later in [4], and h_{ii} is the *i*th diagonal element of the nonlinear "hat matrix" $\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'$ as shown in Gallant (1987) and Rawlings (1988). Studentized residuals are designed to take into account that residuals (r_i) have intrinsically unequal variances even though the theoretical error term (ε_i) is assumed to have constant variance (Draper and Smith, 1981; Montgomery and Peck, 1982; Neter et al., 1990). For a correctly identified function, when the assumptions of the regression analysis are met, the studentized residuals have zero mean and constant variance, and the plot of studentized residuals against the predicted values of the dependent variable will show a homogeneous band. For that reason, the use of the studentized residuals has been recommended by Draper and Smith (1981), Montgomery and Peck (1982), Carroll and Ruppert (1988), and Rawlings (1988).

Figure 3 shows an example of the plot of studentized residuals against the predicted height for the modified logistic function 19 fitted to aspen data with unweighted nonlinear least squares. The plot reveals an obvious unequal error variance problem and suggests that a weighting factor in the form of $w_i = 1/DBH_i^k$ should achieve the desired equality of error variance. This function was then fitted with WLS using six alternative values for k (k= 0.5, 1.0, 1.5, 2.0, 2.5, 3.0). Comparison of the plots of studentized residuals showed that the most appropriate value to be k = 1.0, which resulted in an approximately homogeneous band of the error variance as shown in Figure 4. Accordingly this weighting value $w_i = 1/DBH_i$ was used in all remaining analysis. The calculation and the plotting of the studentized residuals is readily available on the SAS software (SAS Institute Inc., 1985).

The studentized residuals based chosen weight $w_i = 1/DBH_i$ was shown appropriate for all the major Alberta species. The weight implies that the variance of H_i or the variance of the errors is proportional to the corresponding DBH_i. The weight also agrees with the weight chosen by Larsen and Hann (1987) and Wang and Hann (1988) for major Oregon tree species and by Farr et al. (1989) for western hemlock (*Tsuga heterophylla* (Raf.) Sarg.) and Sitka spruce (*Picea sitchensis* (Bong.) Carr.) in Alaska and British Columbia based on different procedures.

The fitting of the nonlinear height-diameter functions for various species groups was accomplished using the PROC NLIN procedure on SAS statistics software (SAS Institute Inc., 1985). The Gauss-Newton method as described in Gallant (1987) was applied in all the function fittings. To ensure the solution is the global rather than the local least squares estimates, different initial values of the parameters were chosen for the fits.

Comparison criteria

Three different criteria are selected for judging the performance of the chosen height-diameter functions:

1. *The asymptotic t-statistic*

For any appropriate height-diameter function, the estimated coefficients should have small asymptotic standard errors. The asymptotic t-statistic, which tests the null hypothesis that each coefficient is zero, should be significant, meaning that the coefficient is reasonable to be included in the function. It is relatively common in nonlinear modelling situations that a function may include many unnecessary parameters, which may result in serious statistical consequences and lead to convergence difficulties in parameter estimation. The principle of parsimony which favours the simpler functions with good statistical properties should be preferred and the problem of overparameterization (Draper and Smith, 1981; Ratkowsky, 1990), whenever possible, should be avoided.

2. *Mean squared error (MSE)*

Mean squared error is error sum of squares divided by its degrees-of-freedom:

$$[4] \quad MSE = \frac{SSE(\theta)}{n-p}$$

where SSE(θ) is defined in [2], n is the number of observations, p is the number of parameters included in the function. MSE is an asymptotically unbiased estimate of the error variance which is often used as a measure of how well the function fits the data. Unlike the use of the coefficient of determination (R^2), which always tend to increase as extra parameters are added, the MSE takes account of the number of parameters and may increase as extra parameters are added if the reduction in the error sum of squares is not sufficient to compensate for the loss in the number of error degrees of freedom (Kvålseth, 1985; Freund and Littell, 1986). The use of the other criteria such as those described by Draper and Smith (1981), Judge et al. (1988), and Neter et al. (1990), including the adjusted coefficient of determination R_a^2, which is a rescaling of R^2 by degrees of freedom, and the C_p statistic, is closely related to MSE and will likely lead to the same conclusions (Montgomery and Peck, 1982; Rawlings, 1988).

3. *The plot of studentized residuals against the predicted height*

Various residual plots that portray the discrepancy between the original data and the fitted functions have been regarded as among the most important criteria for judging the appropriateness of the fitted functions. They have been used extensively in detecting model inadequacies and studying the validity of the regression assumptions (Belsley et al. 1980; Draper and Smith, 1981; Cook and Weisberg, 1982; Rawlings, 1988). The plot of studentized residuals against the predicted height is used in this analysis. An adequately fitted height-diameter function should not show a consistent underestimate or overestimate for the dependent variable height. The plot should reveal an approximately homogeneous band of the error variance.

Results and discussion

Table 4 shows the least squares estimates of the parameters. The asymptotic t-statistics for testing the null hypothesis that each parameter is zero are calculated, and the insignificant parameters are marked. The weighted MSE are summarized in Table 5. The R^2 values for the fitted functions ranged from 0.70 to 0.92, with the average about 0.85.

Results in Table 4 indicate that for two parameter functions 1 to 9, there is no problem in parameter estimations. With the exception of the parameter *a* in function 7 for species group 6b, all the t-statistics for the parameters of the functions are significant at $\alpha = 0.05$ level. The weighted MSE results of the two parameter functions shown in Table 5 indicate that functions 3, 4, and 5 have lower MSE values compared the others, with function 4 generally gives most satisfactory results. Function 8 has very poor performance with large MSE values. The examination of the plots of studentized residuals for function 8 showed biased height estimates for all species groups when DBH is small. The performance of the remaining two parameter functions is roughly the same and can be regarded as intermediate.

Judged from the plots of studentized residuals and the weighted MSE values, the three parameter functions 10 to 19 generally perform better than the two parameter functions. Parameter *a* in function 16 showed several insignificant t-statistics (Table 4). The parameter estimates for the remaining functions are generally satisfactory, with a few exceptions of insignificant t-statistics in functions 10, 12, 15, and 19 for species group 4a and function 15 parameter *b* for species group 6b. All insignificant t-statistics occur when the sample sizes are relatively small. In terms of the weighted MSE values for three parameter functions 10 to 19 (Table 5), functions 12, 13, 15, 18, and 19 generally give lower values. Functions 10 and 14 give rather similar results and can also be regarded as satisfactory. Function 17 has large MSE values and the plots of studentized residuals showed biased estimates when DBH is small. Occasionally, function 11 fit the data well in some particular cases but perform poorly in general.

Although the four parameter function 20 fitted the data well when the sample size is large (such as for species group 1 and 3), the function failed to converge for species group 2b and 6b, and in fitting for species group 4a, has resulted in insignificant t-statistics for parameters *b*, *c*, and *d* (Table 4). Several additional four parameter functions (include Bailey's (1980) function) fitted but not reported here also suggested that they might perform well for large samples, however, insignificant t-statistics occured frequently, and in many cases, failed to converge or converged at local rather than the global minimum when the sample size was small. The gain of using the four parameter function may not be substantial. Depending on the choice of the initial values of the parameters and the size of the samples, the fittings of the four parameter functions may also be rather time consuming.

In terms of the fit of the functions for each species group, several functions may give similar results and perform nearly equally well. However, judging from the weighted MSE values, the asymptotic t-statistics of the parameters, and the principle of parsimony, the following functions are most appropriate for each species group taken independently of the others:

1). The Chapman-Richard function 12 for white spruce;

2). The fractional function 16 for lodgepole pine, whitebark pine, and limber pine;

3). The Gompertz function 14 for jack pine;

4). The Weibull function 13 and the modified Schnute function 15 for aspen;

5). The two parameter Michaelis-Menten function 3 for white birch;

6). The Mitscherlich function 4 for balsam poplar;

7). The modified exponential function 10 for black spruce and engelmann spruce;

8). The modified logistic-type function 19 for balsam fir;

9). The Gompertz function 14 for douglas fir, alpine fir, alpine larch, tamarack, and western larch.

Conclusions and recommendations

This comparison of nonlinear height-diameter functions shows that depending on the sample sizes and the species groups, many functions perform well in describing the height-diameter relationships for major Alberta species. The choice of a particular function may depend on the relative ease of achieving convergence to a solution, the function's mathematical properties and its biological interpretations, and sometimes personal preference. Although any function may be considered superior or inferior in a particular situation, *in general*, the following functions are recommended for general use since they often give relatively lower MSE values, significant asymptotic t-statistics, and satisfactory plots of studentized residuals against the predicted values of the dependent variable. Any one of these functions could be used when the same model form is desirable for several species. The recommended functions also have the flexibility to assume various shapes with different parameter values and produce satisfactory curves under most circumstances. All the curves assume biologically reasonable shapes that prevent the unrealistic height predictions in the cases of extrapolating the functions beyond the range of the original data.

1). Function 12: $H=1.3+a(1-e^{-bD})^c$. This three parameter Chapman-Richards function has been used extensively in describing the height-age relationships. The results shown in this analysis indicate that the function is also well suited for modelling height-diameter relationships. One limiting form of the function - equation 14 also gives satisfactory fits, especially when the sample size is relatively small, such as the fits for species group 2b, 4a, and 6b. However, equation 14 may not fit as well as either the Weibull-type function or the Chapman-Richards function when the sample size is large. A cautionary note for the Chapman-Richards function is that it approaches the asymptote too quickly when the dependent variable is only weakly related to the independent variable.

2). Function 13: $H=1.3+a(1-e^{-bD^c})$. This Weibull-type function is consistently among the best height-diameter functions. It is interesting to see that in fitting species group 4a data, the three or four parameter Chapman-Richards function fails to produce significant t-statistic for the parameter *b*. However, the Weibull function performs better and gives all significant t-statistics for the parameters.

3). Function 19: $H=1.3+a/(1+b^{-1}D^{-c})$. Although termed as the modified logistic-type function, the function is quite different from the commonly used logistic function (such as equation 11). It has an inflection point at: $H=1.3+a(c-1)/2c$, $D=[b(c+1)/(c-1)]^{-1/c}$, which allows the function to accommodate many shapes that are commonly described by other sigmoidal functions. The function has an asymptote at $H = 1.3 + a$. It fits the height-diameter relationship well and is consistently among the best height-diameter functions. As examples, the fits of the function for white spruce and aspen are shown in Figure 5 and Figure 6. The plot of studentized residuals against the predicted height for aspen was shown in Figure 4. It is clear that the function appropriately fits the data.

4). Function 18: $H=1.3+a\cdot e^{b/(D+c)}$. This exponential-type function is particularly well suited for deciduous species. It has an asymptote at $H=1.3+a$ and an inflection point at $H=1.3+a\cdot e^{-2}$, $D=-c-b/2$. The function might slightly overestimate height for large diameter trees.

5). Function 15: $H=\{y_1^{\,b}+(c^b-y_1^{\,b})[1-e^{-a(D-D_0)}]/[1-e^{-a(D_2-D_0)}]\}^{1/b}$. This modified Schnute function (with origin set at $D = 0$, $H = 1.3$) was shown to fit the height-diameter relationships reasonably well. With the versatility of this function and its abilities to describe various biological shapes, and the relatively easy parameter estimations and interpretations, further application and evaluation of the function should prove useful. The function may not be so complicated as it looks.

It should be straightforward to extend the functions analysized in this study to model other forestry relationships such the volume-age, height-age, and basal area-age functions. The parameter estimates in Table 4, if appropriately scaled, might be used as the initial values in new applications.

Acknowledgements

The research was supported by the Alberta Forest Development Research Trust Fund and the Department of Forest Science, University of Alberta. Data were kindly provided by the Timber Management Branch of the Alberta Forest Service. Special thanks to Dave Morgan, Tom Lakusta, and Brian Maier of the Alberta Forest Service, and Yanning Peng of the Department of Mathematics at the University of Alberta for providing valuable help and cooperation.

References

AFS. 1988. Alberta Phase 3 Forest Inventory: Tree Sectioning Manual. Alberta Forest Service, Edmonton, Alberta. Pub. No. T/168. Revised 1988 (formerly ENF report Dept. 56).

Bailey, R.L. 1980. The potential of Weibull-type functions as a flexible growth curves: Discussion. Can. J. For. Res. 10: 117-118.

Bates, D.M., and Watts, D.G. 1980. Relative curvature measures of nonlinearity. J. Roy. Stat. Soc. B-42: 1-16.

Belsley, D.A., Kuh, E., and Welsch, R.E. 1980. Regression diagnostics: identifying influential data and sources of collinearity. John Wiley & Sons, New York.

Bolstad, P.V., and Allen, H.L. 1987. Height and diameter growth response in loblolly pine stands following fertilization. For. Sci. 33: 644-653.

Bredenkamp, B.V., and Gregoire, T.G. 1988. A forestry application of Schnute's generalized growth function. For. Sci. 34: 790-797.

Buford, M.A. 1986. Height-diameter relationship at age 15 in loblolly pine seed sources. For. Sci. 32: 812-818.

Burk, T.E., and Burkhart, H.E. 1984. Diameter distributions and yields of natural stands of loblolly pine. Sch. of For. and Wildlife Resources, Virginia Polytechnic Inst. & State Univ., Publ. FWS-1-84.

Burkhart, H.E., Parker, R.C., Strub, M.R., and Oderwald, R.G. 1972. Yield of old-field loblolly pine plantations. Sch. of For. and Wildlife Resources, Virginia Polytechnic Inst. & State Univ., Publ. FWS-3-72.

Burkhart, H.E., and Strub, M.R. 1974. A model for simulation of planted loblolly pine stands. In growth models for tree and stand simulation (J. Freies, ed.). Roy. College of For., Stockholm, Sweden, p. 128-135.

Cao, Q.V., Burkhart, H.E., and Lemin, R.C. 1982. Diameter distributions and yields of thinned loblolly pine plantations. Sch. of For. and Wildlife Resources, Virginia Polytechnic Inst. & State Univ., Publ. FWS-1-82.

Carroll, R.J., and Ruppert, D. 1988. Transformation and weighting in regression. Chapman and Hall, London.

Cook, R.D., and Weisberg, S. 1982. Residuals and Influence in Regression. Chapman and Hall, London.

Curtis, R.O. 1967. Height-diameter and height-diameter-age equations for second-growth Douglas-fir. For. Sci. 13: 365-375.

Curtis, R.O., Clendenen, G.W., and DeMars, D.J. 1981. A new stand simulator for coast Douglas-fir: DFSIM user's guide. USDA For. Serv. Gen. Tech. Rep. PNW-128.

Draper, N., and Smith, H. 1981. Applied Regression Analysis. 2nd edition. John Wiley & Sons, New York.

Ek, A.R. 1973. Performance of regression models for tree height estimation with small sample size. Statistics in For. Res., IUFRO Sub-group S6.02. Vancouver, B.C., Canada. pp. 67-80.

Farr, W.A., DeMars, D.J., and Dealy, J.E. 1989. Height and crown width related to diameter for open-grown western hemlock and Sitka spruce. Can. J. For. Res. 19: 1203-1207.

Freund, R.J., and Littell, R.C. 1986. SAS System for Regression. SAS Institute Inc., Cary, NC.

Furnival, G.M. 1961. An index for comparing equations used in constructing volume tables. For. Sci. 7: 337-341.

Gallant, A.R. 1987. Nonlinear Statistical Models. John Wiley & Sons, New York.

Judge, G.G., Hill, R.C., Griffiths, W.E., Lütkepohl, H., and Lee, T.C. 1988. Introduction to the theory and practice of econometrics. 2nd ed. John Wiley & Sons, New York.

Ker, J.W., and Smith, J.H.G. 1955. Advantages of the parabolic expression of height-diameter relationships. For. Chron. 31: 236-246.

Kozak, A., and Yang, R.C. 1978. Height-diameter curves, another application of the Weibull function in forestry. 5th Meet. IUFRO, Sub-group S6.02. Freiburg, German Federal Republic, June 12-17, 1978.

Kvålseth, T.O. 1985. Cautionary note about R^2. The American statistician, 39: 279-286.

Larsen D.R., and Hann, D.W. 1987. Height-diameter equations for seventeen tree species in southwest Oregon. For. Res. Lab, Oregon State Univ., Res. Pap. 49.

Larson, B.C. 1986. Development and growth of even-aged stands of Douglas-fir and grand fir. Can. J. For. Res. 16: 367-372.

Loetsch, F., Zöhrer, F., and Haller, K.E. 1973. Forest inventory. Volume 2. BLV Verlagsgesellschaft mbH, Müchen, Germany.

Medawar, P.B. 1940. The growth, growth energy, and aging of the chicken's heart. Proc. Roy. Soc. of London, B-129: 332-355.

Meyer, H.A. 1940. A mathematical expression for height curves. J. For. 38: 415-420.

Moffat, A.J., Matthews, R.W., and Hall, J.E. 1991. The effects of sewage sludge on growth and foliar and soil chemistry in pole-stage Corsican pine at Ringwood Forest, Dorset, UK. Can. J. For. Res. 21: 902-909.

Montgomery, D.C., and Peck, E.A. 1982. Introduction to linear regression analysis. John Wiley & Sons, New York.

Näslund, M. 1936. The forest research institute's thinning experiment in pine forests. Medd. Stat. Skogsförsöksant., 29: 1-172. Swed.

Neter, J., Wasserman, W., and Kutner, M. 1990. Applied linear statistical models. 3rd ed. Irwin, Homewood, IL.

Pearl, R., and Reed, L.J. 1920. On the rate of growth of the population of the United States since 1790 and its mathematical representation. Proc. Nat. Acad. Sci. 6: 275-288.

Prodan, M. 1968. Forest biometrics. English ed., Pergamon Press, Oxford (German ed., 1961).

Ratkowsky, D. A. 1983. Nonlinear regression modeling - a unified practical approach. Marcel Dekker, Inc., New York.

Ratkowsky, D. A. 1990. Handbook of nonlinear regression. Marcel Dekker, Inc., New York.

Ratkowsky, D.A., and Reedy, T.J. 1986. Choosing near-linear parameters in the four-parameter logistic model for radioligand and related assays. Biometrics 42: 575-582.

Rawlings, J.O. 1988. Applied regression analysis - a research tool. Wadsworth, Inc., Belmont, CA.

Richards, F. J. 1959. A flexible growth function for empirical use. J. Exp. Biol. 10: 290-300.

Robertson, T. B. 1923. The chemical basis of growth and senescence. J. B. Lippincott Co., Philadelphia, PA.

SAS Institute Inc. 1985. SAS user's guide: statistics. Version 5 ed., Cary, NC.

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

Schreuder, H.T., Hafley, W.L., and Bennett, F.A. 1979. Yield prediction for unthinned natural slash pine stands. For. Sci. 25: 25-30.

Seber, G.A.F., and Wild, C.J. 1989. Nonlinear regression. John Wiley & Sons, New York.

Sibbesen, E. 1981. Some new equations to describe phosphate sorption by soils. J. Soil Sci. 32: 67-74.

Staebler, G.R. 1954. Standard computations for permanent sample plots. USDA For. Serv. Puget Sound Res. Centre Advisory Committee, Pacific Northwest For. and Range Exp. Stat., Portland, Oregon.

Stage, A.R. 1975. Prediction of height increment for models of forest growth. USDA For. Serv. Res. Pap. INT-164.

Stoffels, A., and van Soest, J. 1953. The main problems in sample plots. 3. height regression. Ned. Boschb. Tijdschr. 25: 190-199. (English summary in For. Abstr. 15: 77).

Trorey, L.G. 1932. A mathematical method for the construction of diameter height curves based on site. For. Chron. 8: 121-132.

Wang, C.H., and Hann, D.W. 1988. Height-diameter equations for sixteen tree species in the central western Willamette valley of Oregon. For. Res. Lab, Oregon State Univ., Res. Pap. 51.

Watts, S.B. 1983. Forestry handbook for British Columbia. 4th ed., For. Undergrad. Soc., Vancouver, B.C.

Winsor, C.P. 1932. The Gompertz curve as a growth curve. Proc. Nat. Acad. Sci. 18: 1-7.

Wykoff, W.R., Crookston, N.L., and Stage, A.R. 1982. User's guide to the stand prognosis model. USDA For.

Serv. Gen. Tech. Rep. INT-133.

Yang, R.C., Kozak, A., and Smith, J.H.G. 1978. The potential of Weibull-type functions as a flexible growth curves. Can. J. For. Res. 8: 424-431.

Yang, Y.C., and Feng F.L. 1989. The application of Schnute growth function to the analysis of stand structure of man-made forests in Taiwan. Quarterly J. of Chinese Forestry 22: 3-17.

Zakrzewski, W.T., and Bella, I.E. 1988. Two new height models for volume estimation of lodgepole pine stands. Can. J. For. Res. 18: 195-201.

Table 1. Species, species code, and species group

Species group	Species	Species code	Scientific name
1	White spruce	SW	*Picea glauca* (Moench) Voss
2a	Lodgepole pine	PL	*Pinus contorta* var. *latifolia* Engelm.
	Whitebark pine	PW	*Pinus albicaulis* Engelm.
	Limber pine	PF	*Pinus flexilis* James
2b	Jack pine	PJ	*Pinus banksiana* Lamb.
3	Aspen	AW	*Populus tremuloides* Michx.
4a	White birch	BW	*Betula papyrifera* Marsh.
4b	Balsam poplar	PB	*Populus balsamifera* L.
5	Black spruce	SB	*Picea mariana* (Mill.) B.S.P.
	Engelmann spruce	SE	*Picea engelmannii* Parry
6a	Balsam fir	FB	*Abies balsamea* (L.) Mill.
6b	Douglas fir	FD	*Pseudotsuga menziesii* (Mirb.) Franco.
	Alpine fir	FA	*Abies lasiocarpa* (Hook.) Nutt.
	Alpine larch	LA	*Larix lyallii* Parl.
	Tamarack	LT	*Larix laricina* (Du Roi) K. Koch
	Western larch	LW	*Larix occidentalis* Nutt.

Table 2. Species group based tree summary statistics

Species group	Number of sample trees	DBH (cm)				Height (m)			
		Mean	Min	Max	Std	Mean	Min	Max	Std
1	3101	26.41	1.20	89.00	12.19	20.09	1.70	38.40	6.98
2a	3199	22.10	1.10	66.60	8.59	18.11	1.72	37.60	5.18
2b	659	18.01	1.60	45.00	9.81	14.74	2.58	28.20	6.38
3	3647	21.36	1.10	64.40	10.12	18.77	2.23	31.94	5.46
4a	102	12.11	1.60	32.00	5.87	11.88	3.18	21.50	4.13
4b	510	22.75	1.10	52.90	9.79	17.76	2.90	31.95	4.88
5	1628	14.10	1.10	55.30	6.08	12.20	1.76	30.63	4.26
6a	508	21.15	1.30	53.00	9.19	16.11	1.78	31.40	5.50
6b	135	20.60	3.30	48.70	9.72	13.26	3.35	22.33	4.98

Table 3. Nonlinear height-diameter functions selected for comparison

Number and form[1]	References
1. $H=1.3+aD^{b}$	Stoffels and Van Soest, 1953; Stage, 1975; Schreuder et al., 1979
2. $H=1.3+e^{a+b/(D+1)}$	Wykoff et al., 1982
3. $H=1.3+aD/(b+D)$	Bates and Watts, 1980; Ratkowsky, 1990
4. $H=1.3+a(1-e^{-bD})$	Meyer, 1940; Farr et al., 1989; Moffat et al., 1991
5. $H=1.3+D^{2}/(a+bD)^{2}$	Näslund, 1936; Loetsch, et al., 1973
6. $H=1.3+a\cdot e^{b/D}$	Ek, 1973; Burkhart and Strub, 1974; Cao et al., 1982; Burk and Burkhart, 1984; Buford, 1986; Bolstad and Allen, 1987; Zakrzewski and Bella, 1988
7. $H=1.3+10^{a}D^{b}$	Larson, 1986
8. $H=1.3+aD/(D+1)+bD$	Watts, 1983
9. $H=1.3+a[D/(1+D)]^{b}$	Curtis, 1967; Pordan, 1968
10. $H=1.3+e^{a+bD^{c}}$	Curtis et al., 1981; Larsen and Hann, 1987; Wang and Hann, 1988
11. $H=1.3+a/(1+b\cdot e^{-cD})$	Pearl and Reed, 1920; Robertson, 1923
12. $H=1.3+a(1-e^{-bD})^{c}$	Richards, 1959
13. $H=1.3+a(1-e^{-bD^{c}})$	Kozak and Yang, 1978; Yang et al., 1978
14. $H=1.3+a\cdot e^{-b\cdot e^{-cD}}$	Winsor, 1932; Medawar, 1940
15. $H=\{y_1^{b}+(c^{b}-y_1^{b})[1-e^{-a(D-D_0)}]/[1-e^{-a(D_2-D_0)}]\}^{1/b}$	Schnute, 1981; Bredenkamp and Gregoire, 1988; Yang and Feng, 1989
16. $H=1.3+D^{2}/(a+bD+cD^{2})$	Curtis, 1967; Pordan, 1968
17. $H=1.3+aD^{bD^{-c}}$	Sibbesen, 1981
18. $H=1.3+a\cdot e^{b/(D+c)}$	Ratkowsky, 1990
19. $H=1.3+a/(1+b^{-1}D^{-c})$	Ratkowsky and Reedy, 1986
20. $H=1.3+a(1-b\cdot e^{-cD})^{d}$	Richards, 1959

[1] H=total tree height in metres; D=DBH in centimeters; a, b, c, d=parameters to be estimated; e=base of the natural logarithm (≈2.71828); 1.3=a constant commonly used to avoid the prediction of a height less than 1.3 metres when DBH is small. For equation 15: $y_1=1.3$, $D_0=0.0$, $D_2=100.0$.

Table 4. Comparison of nonlinear height-diameter functions: parameter estimations

Function	Parameter	Estimates for various species groups								
		1	2a	2b	3	4a	4b	5	6a	6b
1	a	1.7313	2.0196	1.3150	2.8211	1.9024	2.6947	1.2137	1.2469	1.1000
	b	0.7353	0.6899	0.8126	0.6056	0.6986	0.5871	0.8344	0.8163	0.7954
2	a	3.6042	3.4766	3.3789	3.3910	3.0097	3.3238	3.2087	3.4184	3.2256
	b	-16.1901	-13.8574	-12.6489	-10.1272	-7.5330	-10.9470	-11.3747	-14.3731	-14.1907
3	a	62.9784	51.4152	65.6462	39.9983	33.4618	37.0257	59.4777	58.3695	51.2611
	b	58.0915	43.2873	65.5679	24.7274	24.2608	26.0386	60.7484	59.0756	64.0364
4	a	38.8548	32.4692	37.9810	27.1294	21.3657	25.3302	34.1127	34.1281	29.9225
	b	0.0270	0.0349	0.0260	0.0549	0.0614	0.0512	0.0283	0.0285	0.0263
5	a	1.8737	1.6413	1.6840	1.1800	1.0601	1.3209	1.5986	1.8069	2.0261
	b	0.1519	0.1639	0.1666	0.1753	0.2089	0.1813	0.1814	0.1663	0.1805
6	a	35.2854	30.8991	27.5419	28.2674	18.3182	26.6049	22.7872	29.3762	23.8673
	b	-14.4531	-12.1948	-10.7183	-8.5907	-5.6927	-9.4854	-9.3829	-12.8412	-12.3567
7	a	0.2388	0.3048	0.1189	0.4509	0.2793	0.4305	0.0838	0.0953	0.0413*
	b	0.7350	0.6903	0.8126	0.6053	0.6986	0.5871	0.8347	0.8167	0.7955
8	a	3.8180	4.9317	2.0670	6.4194	3.2636	6.5487	1.5058	1.3123	1.8794
	b	0.5738	0.5487	0.6401	0.5349	0.6306	0.4507	0.6746	0.6418	0.4951
9	a	35.9867	31.6026	28.3882	28.9552	19.2299	27.1752	23.6995	29.9060	24.4681
	b	15.2897	13.0009	11.6357	9.3290	6.5500	10.1979	10.3221	13.5674	13.2207
10	a	4.3207	4.2512	6.1440	3.8984	6.1541	4.3133	4.6202	4.0034	4.4488
	b	-6.5426	-5.7514	-6.6024	-4.7580	-5.8482	-4.5425	-5.6452	-6.4430	-6.0225
	c	-0.4872	-0.4588	-0.2204	-0.5182	-0.1778*	-0.3614	-0.3577	-0.5375	-0.3793
11	a	26.0850	23.7434	21.8863	22.5297	16.9311	21.5241	17.0593	19.2315	17.3308
	b	8.5482	5.9593	8.5656	5.9461	5.7035	5.0012	8.5954	15.9742	9.7975
	c	0.1339	0.1311	0.1612	0.1704	0.1996	0.1404	0.2063	0.2204	0.1703
12	a	32.0363	29.4214	31.7252	25.7461	25.3245	26.0462	25.0216	23.6894	22.3239
	b	0.0456	0.0457	0.0376	0.0669	0.0409*	0.0464	0.0518	0.0724	0.0522
	c	1.2974	1.1381	1.1150	1.1308	0.8779	0.9465	1.2004	1.6232	1.3270
13	a	31.0481	29.0401	29.8908	25.4088	26.2522	26.1321	24.5127	22.4771	20.8982
	b	0.0209	0.0318	0.0269	0.0486	0.0579	0.0535	0.0308	0.0179	0.0219
	c	1.1973	1.0902	1.1061	1.0892	0.9017	0.9659	1.1361	1.3905	1.2490
14	a	27.8725	25.2831	24.1320	23.5467	18.4726	22.6368	18.8367	20.9530	19.0959
	b	2.8490	2.4343	2.7151	2.3800	2.2367	2.1570	2.8446	3.6061	2.9034
	c	0.0848	0.0873	0.0943	0.1152	0.1235	0.0951	0.1247	0.1259	0.0988
15	a	0.0494	0.0466	0.0450	0.0696	0.0382*	0.0464	0.0536	0.0929	0.0685
	b	0.6387	0.8289	0.7717	0.8151	1.2179	1.0716	0.7411	0.2072	0.4335*
	c	32.4840	30.3314	30.5534	26.8357	26.5976	27.0745	26.1924	23.8101	21.9696
16	a	2.6944	1.4431	0.3504*	0.8408	-0.2324*	0.0038*	1.2706	4.4024	2.4627*
	b	0.6514	0.6806	0.9442	0.4951	0.7813	0.7027	0.8044	0.4670	0.9370
	c	0.0214	0.0233	0.0168	0.0284	0.0273	0.0270	0.0246	0.0311	0.0273
17	a	36.8921	28.4645	39.5300	26.1702	22.7752	22.9433	20.8584	27.8154	35.2386
	b	-13.0405	-16.5206	-8.3474	-13.1935	-7.4274	-20.9985	-14.1796	-15.7403	-8.1497
	c	1.3051	1.5168	1.1040	1.5795	1.3156	1.7680	1.5780	1.4637	1.0545
18	a	43.4552	38.6721	43.7438	33.6553	31.0846	33.2971	31.7946	34.2258	33.0533
	b	-24.1871	-21.4197	-28.1548	-14.5592	-18.4473	-18.4014	-18.5302	-18.7186	-25.2112
	c	5.0167	5.0827	7.3227	3.5766	5.8302	5.5088	4.0490	3.1265	5.8787
19	a	39.3710	37.5445	46.1750	31.3194	41.9635*	34.4682	32.8728	27.6307	28.4451
	b	0.0130	0.0203	0.0174	0.0328	0.0365	0.0369	0.0204	0.0109	0.0146
	c	1.3408	1.2169	1.1253	1.2487	0.9155	1.0589	1.2307	1.5829	1.3299
20	a	32.5525	30.8722	24.4874**	25.4676	20.7813	25.2716	31.3035	23.3678	17.8206**
	b	1.0200	1.0413	0.2528**	0.9687	0.8247*	0.9574	1.0334	0.9716	0.0314**
	c	0.0428	0.0383	0.0878**	0.0709	0.0767*	0.0530	0.0298	0.0766	0.1069**
	d	1.2034	0.9570	9.4899**	1.2419	1.5810*	1.1025	0.8964	1.7781	90.9244**

Note: * - the asymptotic t-statistic for the parameter is not significant at $\alpha = 0.05$ level; ** - convergence is not obtained.

Table 5. Comparison of nonlinear height-diameter functions: weighted mean squared errors

Function	Weighted MSE for various species groups[A]								
	1	2a	2b	3	4a	4b	5	6a	6b
1	0.5082	0.3886	0.2938	0.3863	0.3326	0.3677	0.2685	0.3366	0.3359
2	0.4675	0.3800	0.3564	0.3370	0.3781	0.3687	0.2702	0.2577	0.3373
3	0.4596	0.3702	0.2770	0.3265	0.3257(1)	0.3465	0.2571	0.2865	0.3171
4	0.4539	0.3686	0.2748	0.3189(5)	0.3261(2)	0.3454(1)	0.2566	0.2813	0.3147
5	0.4443	0.3685	0.2976	0.3218	0.3557	0.3542	0.2547	0.2466	0.3124
6	0.4832	0.3891	0.4049	0.3599	0.4268	0.3828	0.2876	0.2701	0.3584
7	0.5082	0.3886	0.2938	0.3863	0.3326	0.3677	0.2685	0.3366	0.3359*
8	0.5841	0.4210	0.3170	0.4778	0.3539	0.4112	0.2843	0.3894	0.3588
9	0.4751	0.3842	0.3801	0.3477	0.4009	0.3756	0.2784	0.2642	0.3474
10	0.4459	0.3668(2)	0.2804	0.3234	0.3295*	0.3500	0.2502(1)	0.2516	0.3158
11	0.4935	0.3877	0.2701(2)	0.3400	0.3314	0.3598	0.2915	0.2865	0.3035(2)
12	0.4424(2)	0.3674(4)	0.2729	0.3166(2)	0.3268*	0.3458(2)	0.2534(5)	0.2449(2)	0.3089(4)
13	0.4426(3)	0.3675(5)	0.2723(5)	0.3165(1)	0.3271(4)	0.3459(3)	0.2539	0.2458(5)	0.3080(3)
14	0.4597	0.3761	0.2648(1)	0.3236	0.3267(3)	0.3504	0.2694	0.2533	0.3024(1)
15	0.4430(5)	0.3676	0.2717(4)	0.3165(1)	0.3272*	0.3459(3)	0.2544	0.2463	0.3069*
16	0.4435	0.3667(1)	0.2767*	0.3200	0.3275*	0.3472*	0.2510(3)	0.2460	0.3131*
17	0.4658	0.4033	0.2857	0.3795	0.3596	0.4192	0.3014	0.2568	0.3089(4)
18	0.4433	0.3677	0.2708(3)	0.3181(4)	0.3273(5)	0.3463(5)	0.2540	0.2455(4)	0.3090(5)
19	0.4427(4)	0.3671(3)	0.2747	0.3180(3)	0.3279*	0.3469	0.2525(4)	0.2445(1)	0.3100
20	0.4423(1)	0.3668(2)	0.2662*	0.3165(1)	0.3288*	0.3462(4)	0.2507(2)	0.2453(3)	0.3100*

[A] - the smallest five MSE values for each species group with ranks 1 (smallest) to 5 in parentheses; * - the MSE values are not compared because the function has insignificant t-statistic(s) or the convergence is not obtained.

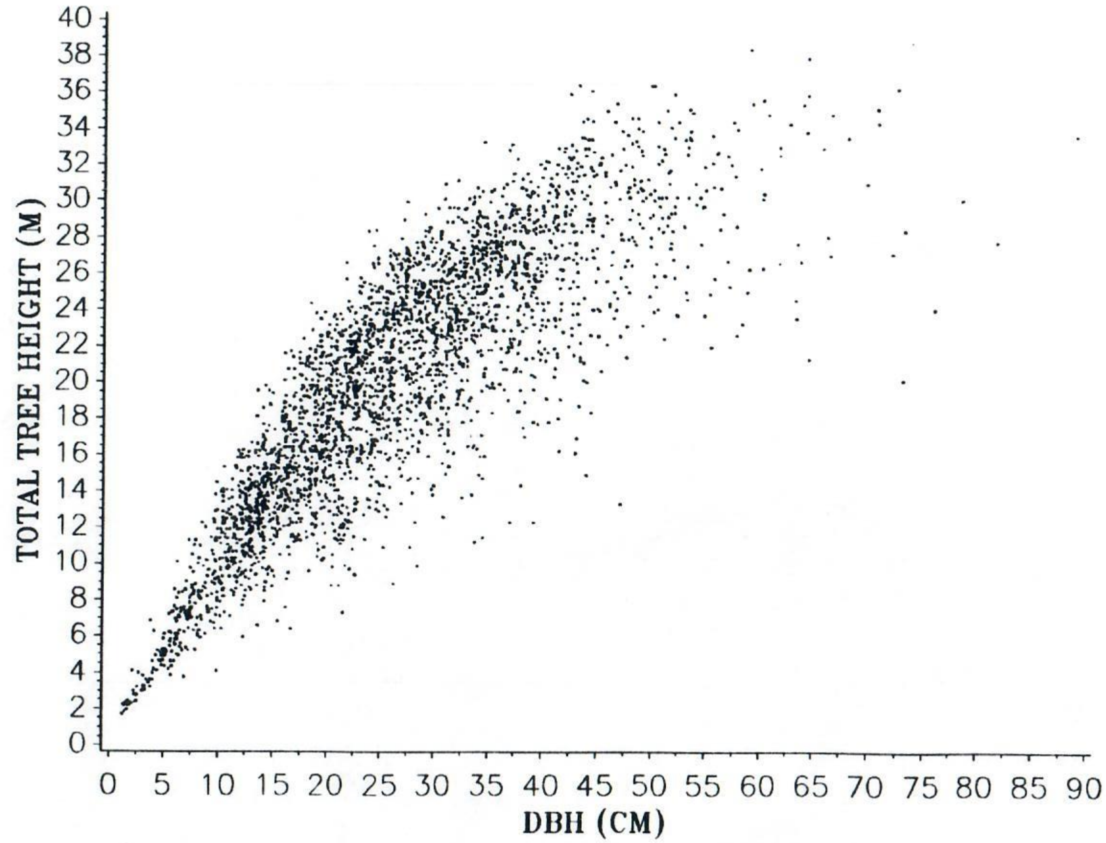

Figure 1. Total tree height against DBH for white spruce (*Picea glauca* (Moench) Voss).

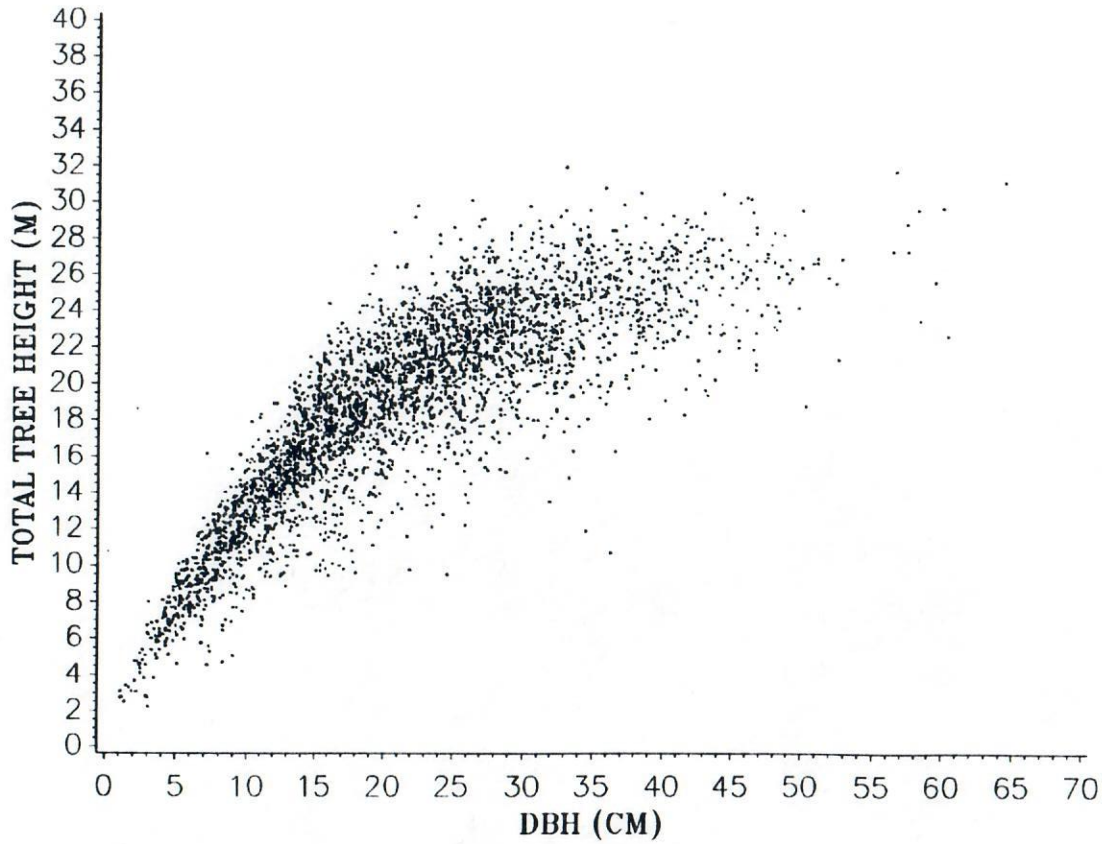

Figure 2. Total tree height against DBH for aspen (*Populus tremuloides* Michx.).

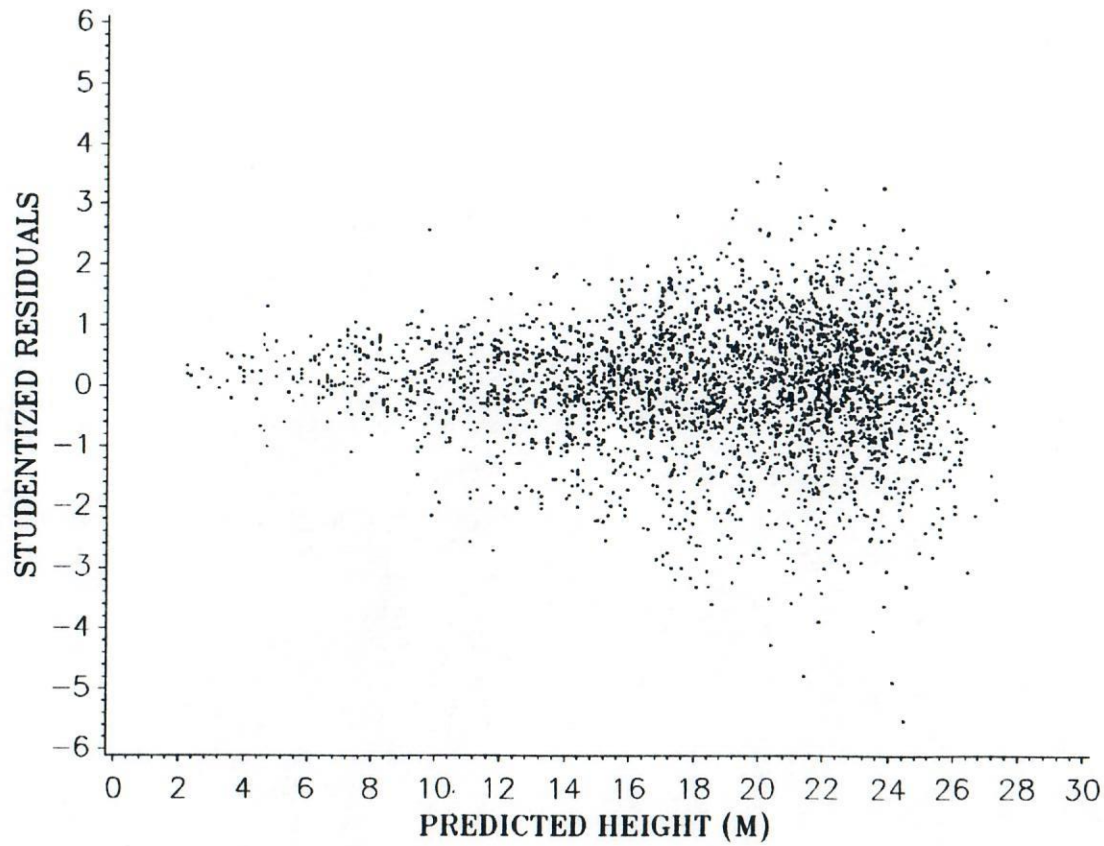

Figure 3. The plot of studentized residuals against the predicted height for aspen (*Populus tremuloides* Michx.). Studentized residuals are obtained by fitting function 19 without weighting.

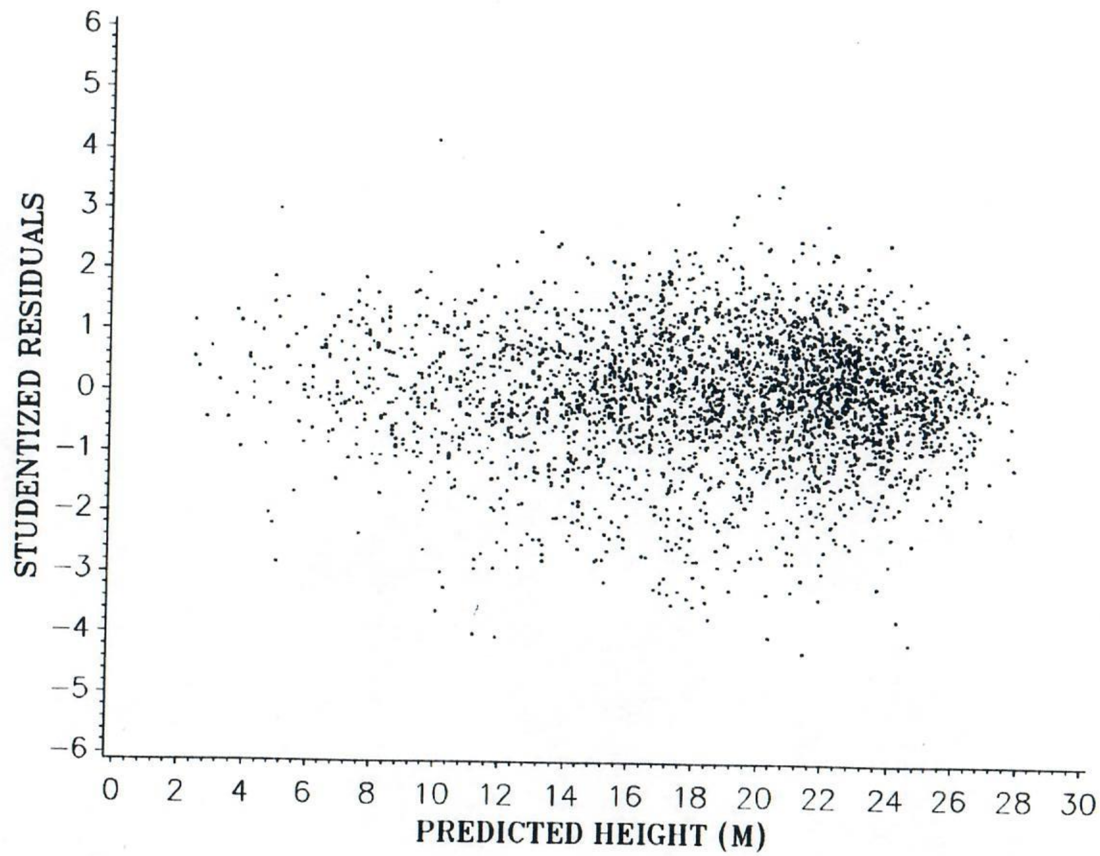

Figure 4. The plot of studentized residuals against the predicted height for aspen (*Populus tremuloides* Michx.). Studentized residuals are obtained by fitting function 19 with weight $w_i = 1/DBH_i$.

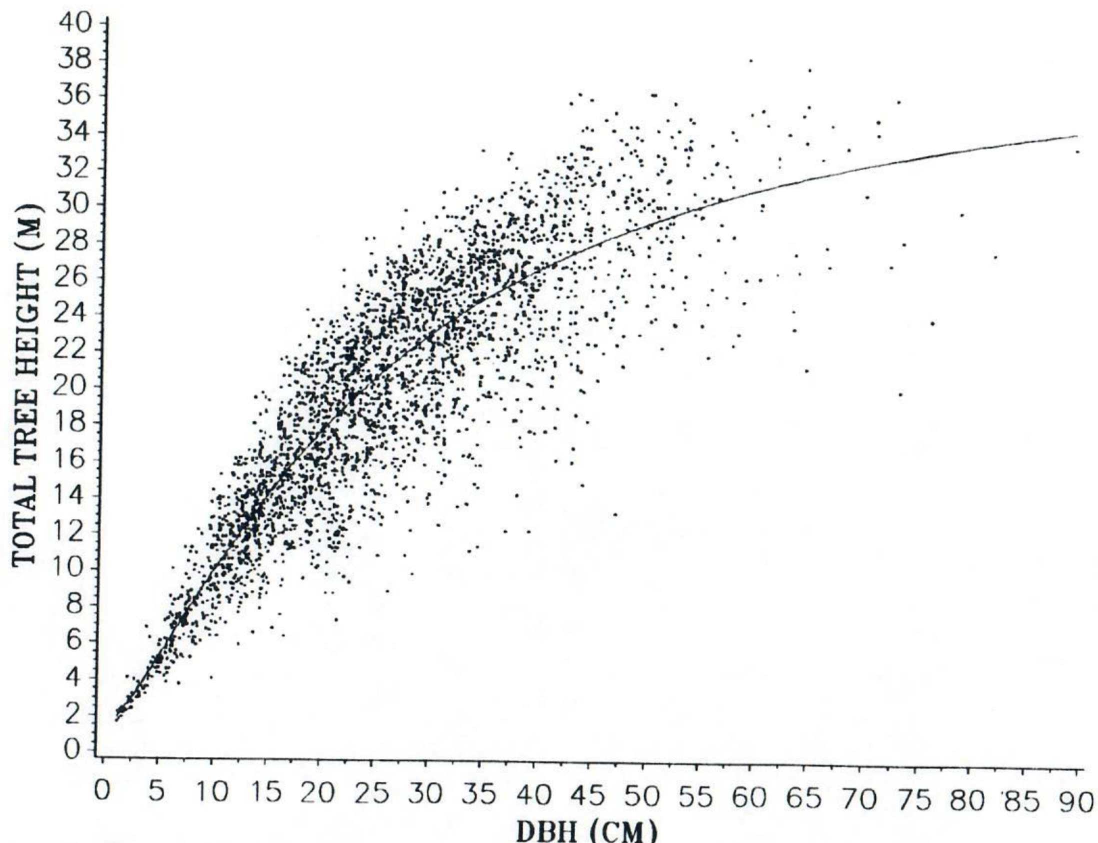

Figure 5. Plot of the original data and the fitted curve for white spruce (*Picea glauca* (Moench) Voss). Curve produced by $H=1.3+39.3710/[1+1/(0.0130\times DBH^{1.3408})]$.

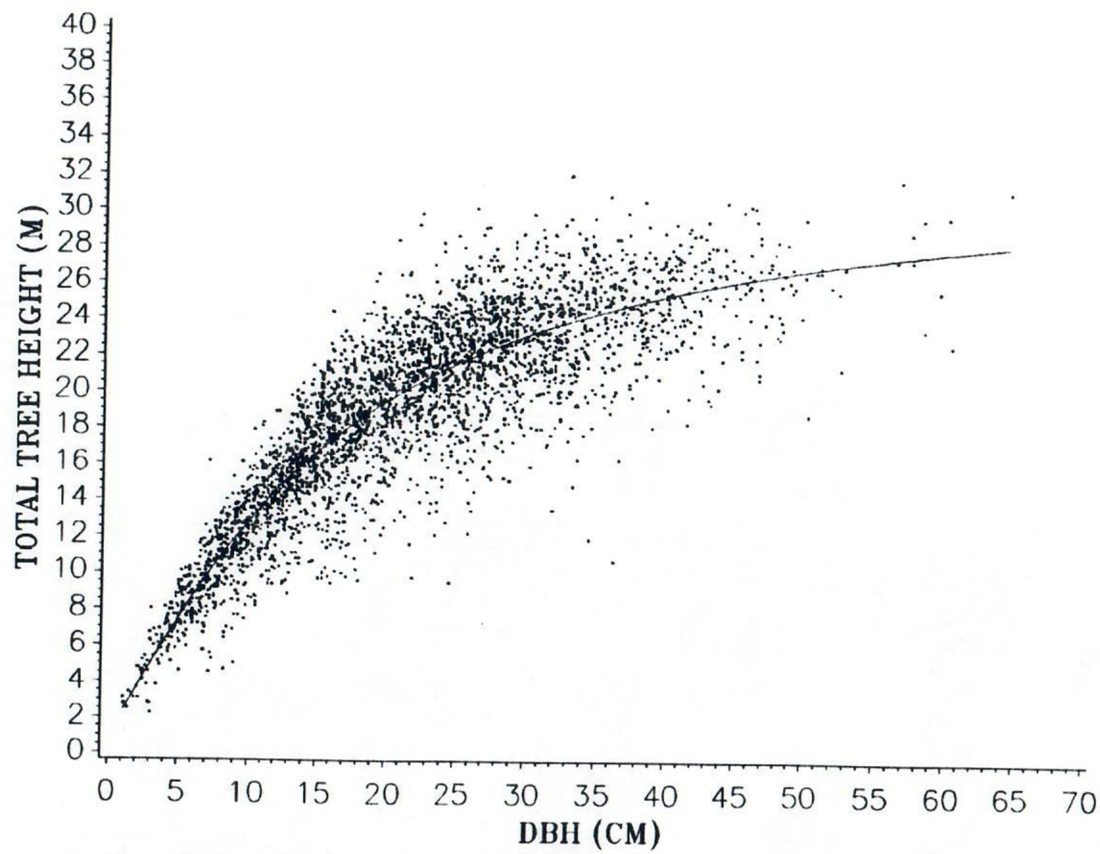

Figure 6. Plot of the original data and the fitted curve for aspen (*Populus tremuloides* Michx.). Curve produced by $H=1.3+31.3194/[1+1/(0.0328\times DBH^{1.2487})]$.