A NATIONAL SYSTEM OF EQUATIONS FOR ESTIMATING OVENDRY MASS OF TREMBLING ASPEN *POPULUS TREMULOIDES* MICHX.

Information Report PI-X-24

F. Evert

Petawawa National Forestry Institute
Canadian Forestry Service
Environment Canada
1983

Catalogue No. Fo46-11/24-1983E
ISSN 0706-1854
ISBN 0-662-12756-0

Additional copies of this publication can be obtained from:

Technical Information and Distribution Centre
Petawawa National Forestry Institute
Canadian Forestry Service
Environment Canada
Chalk River, Ont.
K0J 1J0

Tel.: 613 589-2880

Cette publication est aussi disponible en français sous le titre Système national d'équations pour évaluer la masse anhydre du peuplier faux-tremble Populus Tremuloides Michx.

Contents

A NATIONAL SYSTEM OF EQUATIONS FOR ESTIMATING OVENDRY MASS OF TREMBLING ASPEN POPULUS TREMULOIDES MICHX.

Abstract

This report presents a single national system of equations for estimating the aboveground ovendry mass of single trees and of their individual components for trembling aspen (Populus tremuloides Michx.). This system is based on data from six geographic regions across Canada. When applied to the sample data from individual geographic regions, estimates of aggregate ovendry mass of all sample trees in any of the six regions differed from observed values by not more than 6%, with approximately half the estimates being within 2% of observed values.

Résumé

La rapport présente un système national unique d'équations pour évaluer, chez un peuplier faux-tremble, Populus tremuloides Michx.) la masse anhydre de l'ensemble de sa partie aérienne et de ses éléments. Ce système est basé sur les données provenant de six régions géographiques du Canada. Comparées aux données de chaque région, les estimations de l'aggrégat de la masse anhydre de tout l'échantillon des arbres de n'importe laquelle des régions ne diffèrent pas de plus de 6 % des valeurs observées, la moitié s'en écartant de 2 %.

INTRODUCTION

Standard or two-entry (dbh and height) equations are fundamental to all regional and/or national estimates of ovendry mass for whole trees and for their individual components. They are also fundamental to forecasting growth and yield of ovendry mass because all techniques of forecasting involve, as the first step, estimation of initial values or the taking of an inventory.

There are now at least ten regional systems of standard equations in existence in Canada for estimating the ovendry mass of trembling aspen (Populus tremuloides Michx.), based on direct measurements on at least 700 sample trees. Each regional system consists of equations for estimating ovendry mass of whole trees aboveground, and of each of their components--stem wood, stem bark, live branches, and twigs plus leaves. Some regional systems are based on fewer than 100 sample trees; others, on more than 200. Some samples form the basis for more than one system of equations.

There is reason to think that a single integrated national system of equations would have both practical and theoretical advantages over the many regional systems presently in use.

First, the combined sample of an integrated national system is more likely to represent a wider range of tree sizes than any of the regional samples. Therefore, a single national system would provide accurate estimates for a wider range of tree sizes in all regions than any of the regional systems. Such a sample would also facilitate regression analysis because the variance of the estimated regression coefficients decreases with increase in the range of independent variables (Schumacher and Chapman 1954). In any case, an adequate sample, either regional or national, for a given species should include the full range of diameter classes of the species, and within each diameter class, the full range of height classes. Such regional samples,

F. Evert is a research scientist at the Petawawa National Forestry Institute.

Manuscript approved for publication: 12 August 1983.

however, are uncommon on account of the great expense involved in obtaining them.

Second, a single national system will facilitate the comparison of estimates of regional inventories and the results of national studies such as the interprovincial forest fertilization program (Weetman et al. 1976), parts of which are being carried out in a number of regions. It will also facilitate the comparison of growth and yield data collected in different regions, the exchange of such data, and if justified, integration of such data into a single national growth and yield system.

The objective of this study is to integrate available data sets for trembling aspen collected in various geographic regions of Canada into a single national system of standard or two-entry equations. This system should meet the following requirements:

1) Provide the means for obtaining accurate ovendry mass estimates for trees of all sizes.
2) Provide for the sum of estimated individual components of the whole tree to be equal to the independent estimate of the whole tree.
3) When applied to the sample data from the individual geographic regions, the estimates of the aggregate ovendry mass (Σom_i) of:

    a) all sample trees within a given region, and
    b) all sample trees within each of the three size classes--small, medium, large--in a given region,

    should be within a predetermined percent of the observed values. Because there is no objective method of determining what these predetermined percent values should be, limits were set subjectively--10 percent for all sample trees, and 15 percent for all sample trees within each of the three size classes.

4) When used as a means for forecasting stand growth, by forecasting the individual component variables in a stand ovendry mass equation:

    a) forecasting of the component variables should be practically feasible, and
    b) the sum of estimated individual components of the stand--the ovendry mass of stem wood, stem bark, live branches, and twigs plus leaves, should be equal to the independent estimate of the ovendry mass of the whole stand.

METHODS

Data

Source data were, firstly, computer printouts of individual tree information from six geographic regions and, secondly, published equations from the six regional studies. Published regression data are shown in Table 1, and the distribution of individual tree information is shown in Table 2 by 2 cm dbh and 3 m height classes, for each of the six regions. Some reduction in the number of sample trees in Table 2 from the number in Table 1 is due to a preliminary screening of Ontario data.

Analysis

Solutions to the requirements that the system provide (i) accurate estimates of ovendry mass for trees of all sizes and (ii) that the sum of estimated individual components of the tree must equal the independent estimate of the whole tree, appear to be interrelated. For instance, least squares equations using the same model and fitted from the same data will be "additive" (Kozak 1970), but experience with tree volume equations has shown that they may be poor estimators of small-tree volumes (Evert 1969). Weighted least squares equations tend to be good estimators of the volumes of trees of all sizes (Evert 1969), but their estimates will not necessarily be additive.

To obtain both "additivity", and "good" estimates of small-tree ovendry mass, it was decided to fit weighted least squares equations to individual components of the trees cumulatively instead of separately.

The cumulative variables involved the ovendry mass in kg of individual trees:

stem wood = om(1)
stem wood + bark = om(2)
stem wood + bark + live branches = om(3), and
stem wood + bark + live branches + twigs and leaves = om(4)

Mass of components other than stem wood could be calculated as the difference between estimates obtained from any two appropriate equations. Mass of stem bark, for instance, would be calculated as the difference between estimated masses from equations for om(2) and om(1).

The method used for weighting the ovendry mass residuals of trees was to divide the appropriate ovendry mass of the tree, om(i), by the product of tree basal area at breast height, g, and the total tree height, h. Thus, the following four quantities were regressed on both dbh and height--om(1)/gh, om(2)/gh, om(3)/gh, and om(4)/gh. Ovendry mass [om(i)] equations would be obtained as a separate step, by multiplying both sides of the equations fitted to om(i)/gh, by the product gh.

The first step in the analysis involved verifying whether or not the least squares equations were poor estimators of ovendry mass of small trees, and whether or not the weighted least squares equations provided accurate estimates for trees of all sizes. All models to be tested, both weighted and non-weighted, were based on the 327 observations available from Ontario, mainly because this sample was particularly well balanced for tree sizes (see Table 2). All models involved the fitting of equations to stem wood plus bark because the Petawawa National Forestry Institute (PNFI) did not have stem wood data available separately.

The second step in the analysis involved verifying whether or not equations based on data from a region could be applied in other regions. Standard or two-entry ovendry mass equations should give accurate estimates throughout the range of the species because, although they are rarely applied directly, they will be the principal means of obtaining local ovendry mass equations that will be applied directly. Preparation of local equations from standard equations simply requires (i) construction of a curve of height on dbh for the site or stand to be estimated, and (ii) substitution of height in the standard equation with its estimated value in terms of diameter. Thus, once prepared, standard equations will usually serve as the basis for all local equations. This verification was also based on equations for estimating ovendry mass of stem wood plus bark as was the case in the first step in the analysis.

The third and last step in the analysis involved (i) integrating and regressing all sample data from six regional sources into a single national system of equations, and (ii) verifying fit of the national equations for estimating ovendry mass of trembling aspen throughout its range as far as possible.

All calculations were preceded by a further screening of basic data. Thus, the final calculations involved a total of 675 sample trees instead of 695 trees shown in Table 2.

Because PNFI data did not include information on stem wood except in combination with stem bark, this information was estimated from the stem wood plus bark samples of the PNFI data, using the stem wood/(stem wood+bark) relationship developed from the remaining data. Thus, all four equations in the system including that for stem wood were based on the same 675 sample trees.

RESULTS

Least squares versus weighted least squares equations

Two least squares and four weighted least squares equations are presented in Table 3, and a verification of whether or not they will fit the complete range of basic data from which they were prepared is shown in Figures 1 and 2. The goodness of fit of least squares versus weighted least squares equations cannot be compared using either R^2 or SEE because of the different scale of the residuals involved (see also Appendix 2).

The fit of these equations was verified by expressing their estimates as a percent of observed value and by comparing percentages obtained by the least squares versus the weighted least squares equations. Equation (6) shown in Table 3 was not included in the comparison because, despite an added variable, it is not an improvement over Equation (5). The sum of ovendry mass of basic data estimated by weighted least squares equations does not necessarily equal the sum of their observed values as is the case with estimates of non-weighted equations. The procedure used for correcting this potential bias was (i) to determine the ratio of the two sums, estimated over observed, and (ii) to use the reciprocal of this ratio as a correction factor for the appropriate equation as shown in Figure 2.

The position of the plotted points in Figures 1 and 2 in relation to the 100% line indicates that both least squares and weighted least squares equations provide about equivalent estimates for trees 5 cm and larger in dbh, but for smaller trees, least squares equations are indeed poor estimators of ovendry mass even though the basic sample included a large number of such trees (see Table 2). Weighted least squares equations, however, do provide accurate estimates for trees of all sizes.

Applicability of regional equations throughout the range of the species

Table 4 presents a matrix of values that shows the accuracy of five regional two-entry equations for estimating ovendry mass of aspen stem wood plus bark for six regional samples from across Canada. Each row shows the fit of the five equations when applied to a given regional sample, and each column, the fit of a given regional equation when applied to the six regional samples. All estimates are expressed as a percent of the observed values. Several observations relevant to sampling for and preparation of ovendry mass equations are evident from it. They are:

1) that except for Eq. (5) estimates for samples from Alberta, Ontario (1) and Ontario (2), and Eq. (1) and (4) estimates for the sample from PNFI, all five regional equations provide estimates of aggregate ovendry mass of all trees of the six regional samples that are within ten percent of the observed values. The five regional equations are based on as few as 46 and as many as 252 sample trees.

The failure of Eq. (5) to provide accurate estimates for the regional samples in Alberta, Ontario (1) and Ontario (2) may result from its small basic sample size—only 46 trees, limited range of heights sampled within each diameter class (see Table 2), and the use of a logarithmic regression model:

$$\ln om = -3.0560 + 2.3536 \ln d + 0.2326 \ln h$$

or

$$om = 0.04708 d^{2.3536} h^{0.2326}$$

This model shows that, for a given diameter, ovendry mass increases directly with height to the power of 0.2326. Most plottings show that for a given diameter, ovendry mass increases directly with height to the power of 1.

2) the failure of the Alberta regional equation, Eq. (1), to provide accurate estimates for the PNFI sample that consists of small trees only seems to verify the previously expressed opinion that least squares equations are poor estimators of both small-tree volumes (Evert 1969) and their ovendry mass (see Fig. 1).
3) that Eq. (2) and (3) estimates for the sample from PNFI are within ten percent of the observed values for the aggregate ovendry mass of all trees. But they underestimate the ovendry mass of 0.1-2.0 cm trees, the aggregate estimates being 68.3 and 66.5 percent of the observed values respectively. The model used for Eq. (2) and (3) is:

$$om(i) = bd^2h$$

or

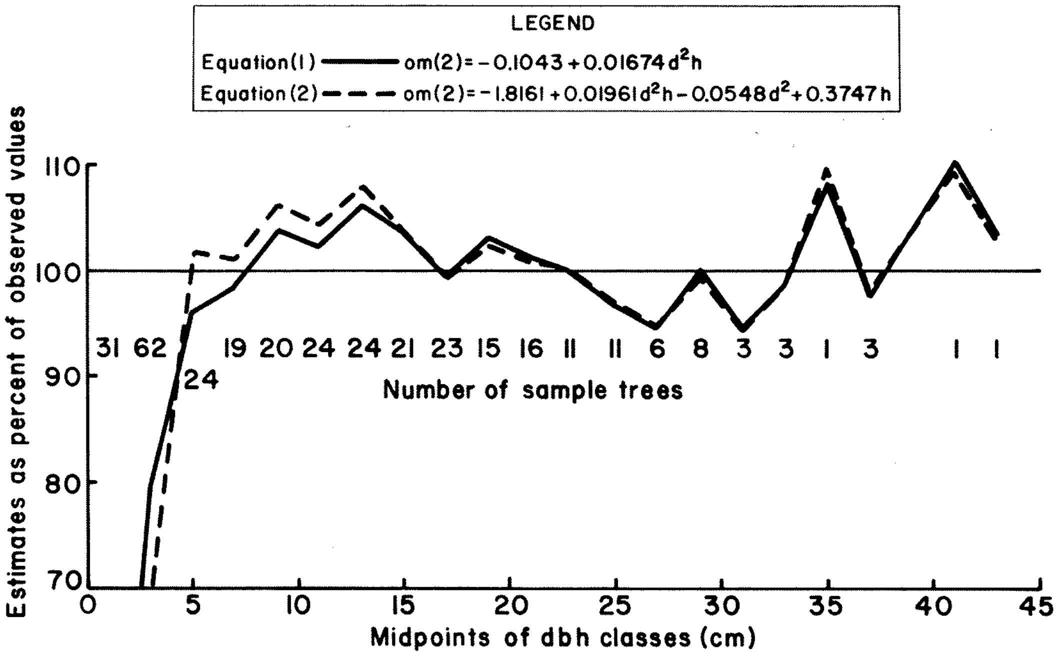

Figure 1. The fit of two least squares equations for estimating the ovendry mass for trees of all sizes.

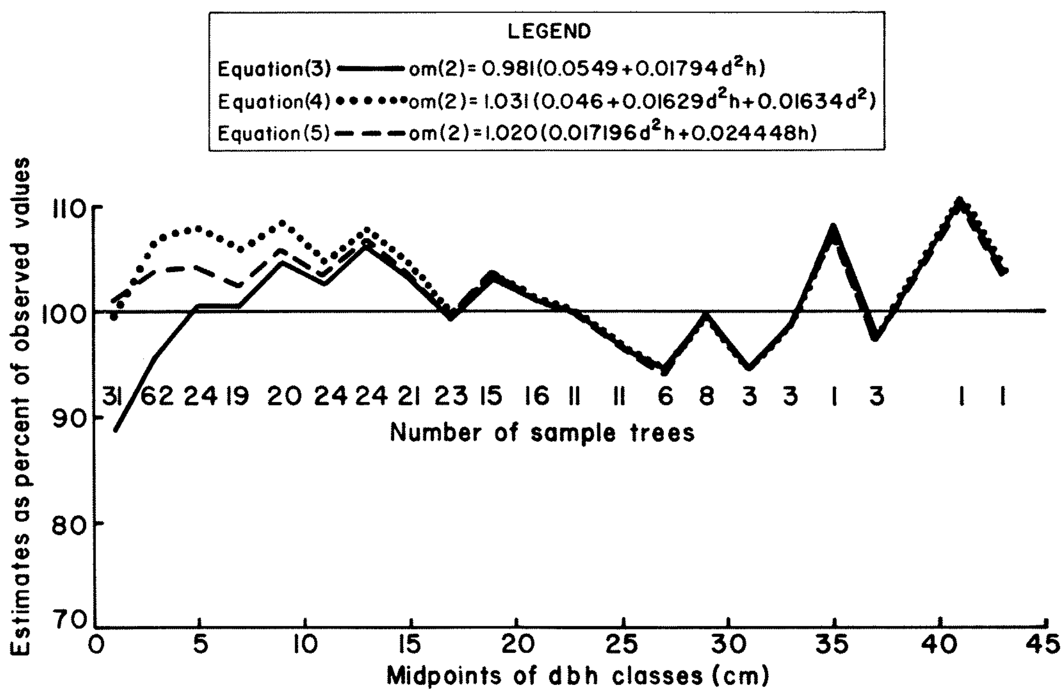

Figure 2. The fit of three weighted least squares equations for estimating the ovendry mass for trees of all sizes.

$$om(i)/d^2h = b$$

In other words, the quantity om(i)/gh is assumed to be independent of tree size. Table 3, however, presents evidence that the coefficient of determination, r^2 or R^2, between the quantity om(i)/gh and tree size can be as high as 0.85 to 0.90.

Preparation and testing of a national system of equations

The national system of equations is presented in Table 5, and a test of its application, in Table 6.

All four equations in Table 5 had the term $1/d^2$ entering first. The term 1/h was also a significant variable in all four equations, but its significance increased consistently from the equation for estimating om(1) to that for estimating om(4). The goodness of fit of the national equations for om(1) and om(4) cannot be compared to that of the corresponding regional equations, by using their respective R^2s and SEEs, because of the different scale of residuals involved (see also Appendix 2).

Table 6 presents a matrix of values that shows the accuracy of ovendry mass estimates obtained with the national system for stem wood, stem bark, live branches, and twigs and leaves, cumulatively.

Each column shows the fit of a given equation when applied to six regional samples. All estimates are again expressed as a percent of the observed values. All estimates of aggregate ovendry mass of all sample trees in any of six regions differ from observed values by no more than six percent; and all aggregate estimates within each of the three size classes in every one of the six regions differ from the observed values by no more than 14 percent.

DISCUSSION AND CONCLUSIONS

It remains for us to elaborate on the evidence presented and to conclude whether or not a single system of standard or two-entry equations would meet the requirements specified in the introduction of this study.

First, does it provide the means of obtaining accurate estimates of ovendry mass for trees of all sizes? The answer to this question seems to depend on the extent of sampling and the method of regression analysis used. Figure 2 provides a strong indication that equations based on a well-distributed sample and fitted by the method of weighted least squares, the weight being the reciprocal of the product of tree basal area and the total tree height, would provide this means.

Second, when applied to the sample data from the individual geographic regions, would the estimates of the aggregate ovendry mass of (a) all sample trees in a given region, and (b) all sample trees within each of the three size classes--small medium, and large--in a given region meet the specified accuracy requirements?

Evidence presented in Table 6 which is summarized below indicates that this requirement will also be met by the single national system of equations, as follows:

Estimated ovendry mass as a percent of observed value	Distribution of cells (from Table 6)			
Class Intervals	Small trees	Medium trees	Large trees	All trees
0.0 - 2.0	12	11	5	11
2.1 - 4.0	4	4	10	6
4.1 - 6.0	4	5	8	7
6.1 - 8.0	2	4	1	-
12.1 - 14.0	2	-	-	-
All	24	24	24	24

Estimates of the aggregate ovendry mass of (a) all sample trees in any of the six regions differ from the observed values by not more than six percent, with approximately half the estimates being within two percent of the observed values, and (b) all sample trees within each of the three size classes in any of the six regions differ from the observed values by not more than 14 percent, with 97 percent of the estimated values being within eight percent of the observed values.

It could be argued that although the national equations meet the specified accuracy requirements, evidence presented in Table 6 suggests that they would systematically underestimate or overestimate for a given region. But what equations would not underestimate or overestimate for a particular sample except on data from which they were derived? What counts is how much and is it acceptable!

Third, when used as a means for forecasting stand growth, will a single national system provide for:

a) the component variables in stand ovendry mass equations that would be practicably predictable, and
b) the sum of estimated individual components of the stand--ovendry mass of stem wood, stem bark, live branches, and twigs and leaves, to be equal to the estimated ovendry mass of the whole stand.

Efficient stand growth forecasting should involve stand ovendry mass equations used with easily predictable stand variables. One system of stand ovendry mass equations will derive

$$om = b_1 d^2 h + b_2 h + b_3 d^2$$

from the basic model used for estimating the ovendry mass of trees (see Table 5). Algebraic conversion of this equation to a stand ovendry mass equation will yield:

$$\begin{aligned} OM &= \Sigma\, om_i \\ &= b_1 \Sigma d_i^2 h_i + b_2 \Sigma h_i + b_3 \Sigma d_i^2 \end{aligned}$$

By factoring $\Sigma d_i^2 h_i$, one obtains:

$$\Sigma d_i^2 h_i = (\Sigma d_i^2)(\Sigma d_i^2 h_i)/(\Sigma d_i^2)$$

and:

$$\begin{aligned} OM = {} & b_1 (\Sigma d_i^2)(\Sigma d_i^2 h_i)/(\Sigma d_i^2) \\ & + b_2 \Sigma h_i + b_3 \Sigma d_i^2 \end{aligned}$$

where OM is stand ovendry mass of stem wood, or stem wood plus bark, etc. for which estimates are being sought,
Σd_i^2 is the sum of squared diameters,
$(\Sigma d_i^2 h_i)/(\Sigma d_i^2)$ is the average height weighted by basal area or Lorey's height (h_L), based on the estimated heights of all trees for which ovendry mass estimates are being sought, and
Σh_i is the sum of estimated heights for all trees for which ovendry mass estimates are being sought.

The sum of estimated heights for all trees could also be expressed as the product of number of trees, N, times the arithmetic mean height of all trees, $\bar{h}$,:

$$\Sigma h_i = N(\Sigma h_i/N) = N\bar{h}$$

to facilitate the forecasting of it when forecasting stand growth. Furthermore, the sum of squared diameters for all trees could also be expressed as the product of number of trees, N, times the quadratic mean diameter, d_g^2,:

$$\Sigma d_i^2 = N d_g^2$$

Thus:

$$OM = N d_g^2 (b_1 h_L + b_3) + b_2 N\bar{h}$$

The answer to the third question, therefore, seems to be affirmative. All the individual component variables in the stand ovendry mass equation as presented above have been used in forecasting stand volume and its growth. It is thus clear that they can also be used to forecast stand ovendry mass and growth.

The method for deriving stand ovendry mass equations algebraically from tree ovendry mass equations will also ensure that the sum of the estimated ovendry mass of individual components of the stand will equal the estimated ovendry mass of the whole stand.

Based on evidence presented, the following conclusions seem to be well supported:

1) To be able to prepare a well-fitting system of ovendry mass equations for both regional and national application, it should be:

    a) based on a sample that includes the full range of diameter classes of the species and, within each diameter class, the full range of height classes. Size range of the sample trees seems to be more important than their geographic distribution;
    b) fitted by the method of weighted least squares instead of the least squares method, the weight being the reciprocal of the product of tree basal area and the total tree height.

2) To ensure additivity of equations when fitted by the method of weighted least squares, fitting should be done to the individual tree components cumulatively instead of separately. Conditions specified by Kozak (1970) would also have to be satisfied. These are:

    a) Exactly the same model should be fitted for all equations.
    b) Any transformation of the dependent variable should be linear in scale.
    c) All equations in a given system should be fitted from the same observations.

3) To maintain the good fit and the additivity of tree ovendry mass equations when used for forecasting stand growth, stand ovendry mass equations should be algebraically derived from tree ovendry mass equations.

As a general recommendation, it seems well worthwhile to examine the feasibility of integrating into a single national system all other regional systems now in use which involve species with a wide geographic range.

ACKNOWLEDGMENTS

Individual tree information used in this study was provided by Canadian Forestry Service Research Centres, in the form of computer printouts. To all those who made their data so readily available, the author expresses his most sincere thanks.

Maritimes FRC	M.F. Ker
Newfoundland FRC	R.S. van Nostrand
Northern FRC	I.E. Bella and T. Singh
Petawawa National Forestry Institute	I.S. Alemdag

Computing was done by Mr. W.A. Kean, and the drafting of the figures by Mr. D.J. McGuire. The author wishes to thank both of them for their assistance.

REFERENCES

Alemdag, I.S.; Horton, K.W. 1981. Single-tree equations for estimating biomass of trembling aspen, largetooth aspen and white birch in Ontario. For. Chron. 57(4):169-173.

Berry, A.B.; Stiell, W.M. 1978. Effect of rotation length on productivity of aspen sucker stands. For. Chron. 54(5):265-267.

Evert, F. 1969. Use of form factor in tree-volume estimation. J. For. 67(2):126-128.

Johnstone, W.D.; Peterson, E.B. 1980. Above-ground component weights in Alberta Populus stands. Can. Dep. Environ., Can. For. Serv., North. For. Res. Cent., Inf. Rep. NOR-X-226. vi+18 p.

Ker, M.F. 1980. Tree biomass equations

for ten major species in Cumberland county, Nova Scotia. Can. Dep. Environ., Can. For. Serv., Maritimes For. Res. Cent., Inf. Rep. M-X-108. iii+26 p.

Kozak, A. 1970. Methods of ensuring additivity of biomass components by regression analysis. For. Chron. 46(5):402-404.

Lavigne, M.B.; van Nostrand, R.J. 1981. Biomass equations for six tree species in central Newfoundland. Can. Dep. Environ., Can. For. Serv., Nfld. For. Res. Cent., Inf. Rep. N-X-199. v+51 p.

Schumacher, F.X.; Chapman, R.A. 1954. Sampling methods in forestry and range management. P. 137, Bulletin 7, revised, School of Forestry, Duke University, Durham, North Carolina. 221 p.

Weetman, G.F.; Krause, H.H.; Koller, E. 1976. Interprovincial forest fertilization program. Results of five-year growth remeasurements in thirty installations: fertilized in 1969, remeasured in 1974. Can. Dep. Environ., Can. For. Serv., For. Techn. Rep. 16. v+34 p.

APPENDIX 1:

TABLES

Table 1. Regression data on the prediction of ovendry mass, om(i), of trembling aspen in six geographic regions

(a) Ontario

(i) Great Lakes - St. Lawrence Forest Region (Alemdag & Horton 1981)
Model om(kg) = bd^2h (cm) (m)
n=128, dbh range 5.2 to 43.5 cm

	b
om(1)[1]	0.014483
om(2)[2]	0.017939
om(3)[3]	0.020610
om(4)[4]	0.021087

(ii) Boreal Forest Region (Alemdag & Horton 1981)
Model om(kg) = bd^2h (cm) (m)
n=96, dbh range from 5.2 to 35.4 cm

	b
om(1)	0.014748
om(2)	0.017496
om(3)	0.019690
om(4)	0.020259

(iii) Petawawa National Forestry Institute (Berry & Stiell 1978)
Model om(kg) = $b_0 + b_1 d^2 h + b_2 (d^2 h)^2$ (cm) (m)
n and dbh range are not reported

	b_0	b_1	b_2
om(3)	0.0992	0.02175	-7.01×10^{-6}

(b) Alberta (Johnstone & Peterson 1980)
Model om(kg) = $b_0 + b_1 d + b_2 d^2 + b_3 d^3 + b_4 h + b_5 d^2 h$ (cm) (m)
n=254, dbh range from 2.0 to 31.5 cm

	b_0	b_1	b_2	b_3	b_4	b_5
om(1)	1.493	0.2384	-0.0046	-0.0004	-0.3040	0.0144
om(2)	1.618	0.3110	0.0178	-0.0005	-0.3916	0.0167
om(3)	-1.367	2.4238	-0.1551	0.0043	-0.8636	0.0183
om(4)	-1.316	2.5077	-0.1566	0.0045	-0.9072	0.0184

(c) Newfoundland (Lavigne & van Nostrand 1981)
Model om(kg) = $b_0 + b_1 d^2 h$ (cm) (m)
n=70, dbh range from 2.2 to 44.7 cm

	b_0	b_1
om(2)	0.2420	0.01691
om(3)	0.1927	0.01933
om(4)	0.5497	0.01987

Table 1. (cont'd)

(d) Nova Scotia (Ker 1980)
Model $\ln om(kg) = b_0 + b_1 \ln d + b_2 \ln h$ (cm)(m)
n=46, dbh range from 1.8 to 33.3 cm

	b_0	b_1	b_2
om(1)	-3.3596	2.3670	0.2548
om(2)	-3.0560	2.3536	0.2326
om(4)	-2.6224	2.4827	-

1/ stem wood

2/ stem wood plus bark

3/ stem wood plus bark plus live branches

4/ stem wood plus bark plus live branches plus twigs and leaves

Table 2. Number of sample trees by location, 2-cm dbh classes, and 3-m height classes

Ont.-G.L.-St. L. For. Reg. (Height range (m))

dbh range (cm)	4.6-7.5	7.6-10.5	10.6-13.5	13.6-16.5	16.6-19.5	19.6-22.5	22.6-25.5	25.6-28.5	All heights
0.1 - 2.0									
2.1 - 4.0									
4.1 - 6.0		5							5
6.1 - 8.0	3	4	4						11
8.1 - 10.0		2	11	2					15
10.1 - 12.0			5	8	1				14
12.1 - 14.0			-	7	7	1			15
14.1 - 16.0			1	1	7	2			11
16.1 - 18.0			-	1	8	3	1		13
18.1 - 20.0			1	-	5	2	-		8
20.1 - 22.0				1	1	2	5		9
22.1 - 24.0						-	5		5
24.1 - 26.0						-	6		6
26.1 - 28.0						1	2		3
28.1 - 30.0							1	1	2
30.1 - 32.0							-	-	-
32.1 - 34.0							1	1	2
34.1 - 36.0							-	-	-
36.1 - 38.0							1	2	3
38.1 - 40.0							-		-
40.1 - 42.0							1		1
42.1 - 44.0							1		1
44.1 - 46.0									
All	3	11	22	20	29	11	24	4	124

Ont.-Bor. For. Reg. (Height range (m))

dbh range (cm)	7.6-10.5	10.6-13.5	13.6-16.5	16.6-19.5	19.6-22.5	22.6-25.5	25.6-28.5	All heights
0.1 - 2.0								
2.1 - 4.0								
4.1 - 6.0	2							2
6.1 - 8.0	5	3						8
8.1 - 10.0	1	3	1					5
10.1 - 12.0	-	2	6	2				10
12.1 - 14.0	1	-	5	3				9
14.1 - 16.0			4	6				10
16.1 - 18.0			1	3	6			10
18.1 - 20.0				4	3			7
20.1 - 22.0				3	4			7
22.1 - 24.0				2	3	1		6
24.1 - 26.0				1	4	-		5
26.1 - 28.0				-	2	1		3
28.1 - 30.0				1	2	3		6
30.1 - 32.0					1	1	1	3
32.1 - 34.0					1		-	1
34.1 - 36.0							1	1
36.1 - 38.0								
38.1 - 40.0								
40.1 - 42.0								
42.1 - 44.0								
44.1 - 46.0								
All	9	8	17	25	26	6	2	93

Ont.-PNFI (Height range (m))

dbh range (cm)	1.6-4.5	4.6-7.5	7.6-10.5	All heights
0.1 - 2.0	28	3		31
2.1 - 4.0	17	45		62
4.1 - 6.0		10	7	17
6.1 - 8.0				
8.1 - 10.0				
10.1 - 12.0				
12.1 - 14.0				
14.1 - 16.0				
16.1 - 18.0				
18.1 - 20.0				
20.1 - 22.0				
22.1 - 24.0				
24.1 - 26.0				
26.1 - 28.0				
28.1 - 30.0				
30.1 - 32.0				
32.1 - 34.0				
34.1 - 36.0				
36.1 - 38.0				
38.1 - 40.0				
40.1 - 42.0				
42.1 - 44.0				
44.1 - 46.0				
All	45	58	7	110

Alta. (Height range (m))

dbh range (cm)	1.6-4.5	4.6-7.5	7.6-10.5	10.6-13.5	13.6-16.5	16.6-19.5	19.6-22.5	22.6-25.5	25.6-28.5	All heights
0.1 - 2.0	1									1
2.1 - 4.0	1	9								10
4.1 - 6.0		9	11							20
6.1 - 8.0		7	9	6						22
8.1 - 10.0		14	3	26	6					49
10.1 - 12.0		12	-	12	8	2	1			35
12.1 - 14.0		4	-	4	13	8	1			30
14.1 - 16.0			-	5	8	5	-	1		19
16.1 - 18.0			1	-	11	1	2	1	1	17
18.1 - 20.0					7	1	-	5	-	13
20.1 - 22.0					7	-	4	-	-	11
22.1 - 24.0					2	1	2	1	1	7
24.1 - 26.0					2	1	-	2	1	6
26.1 - 28.0					1	1	-	2	2	6
28.1 - 30.0						1	2	-	-	3
30.1 - 32.0							1	1	1	3
32.1 - 34.0										
34.1 - 36.0										
36.1 - 38.0										
38.1 - 40.0										
40.1 - 42.0										
42.1 - 44.0										
44.1 - 46.0										
All	2	55	24	53	65	21	13	13	6	252

Table 2. cont.

Nfld.									N.S.							All locations									
1.6-4.5	4.6-7.5	7.6-10.5	10.6-13.5	13.6-16.5	16.6-19.5	19.6-22.5	22.6-25.5	All heights	1.6-4.5	4.6-7.5	7.6-10.5	10.6-13.5	13.6-16.5	16.6-19.5	All heights	1.6-4.5	4.6-7.5	7.6-10.5	10.6-13.5	13.6-16.5	16.6-19.5	19.6-22.5	22.6-25.5	25.6-28.5	All heights
									1						1	30	3								33
2	3							5	3						3	23	57								80
1	3	1						5		3					3	1	32	19							52
	3	-						3		-	3				3		13	21	13						47
		3	1	1				5		1	1				2		15	10	41	10					76
		2	1	-				3		1	-	3			4		13	2	23	22	5	1			66
		1	-	3				4			-	-	1		1		4	2	4	29	18	2			59
			2	3				5			-	1	2	1	4			9	18	19	2	1			49
				3	2			5			-	1	2	-	3			1	1	18	14	11	2	1	48
				3	3	2		8			1	1	1	1	4			1	2	11	14	7	5	-	40
				1	2	-		3				1	1	-	2				1	10	6	10	5	-	32
				2	3	1		6				1	2	3	6				1	6	9	6	7	1	30
				-	4	-		4				-	-	2	2				-	2	8	4	8	1	23
				1	5	-		6				-	-	2	2				-	2	8	3	5	2	20
					2	1		3				1	1	1	3				1	1	5	5	4	1	17
					2	1		3						1	1						3	3	2	2	10
								-						2	2						2	1	1	1	5
								-															-	1	1
								-															1	2	3
								-															-		-
							1	1															2		2
							-	-															1		1
							1	1															1		1
3	9	7	4	17	23	5	2	70	4	5	5	9	10	13	46	54	137	65	105	130	94	54	44	12	695

Table 3. Equations for estimating ovendry mass, based on 327 observations from Ontario

Equation No.	Equation	r^2 or R^2	SEE %
	Least squares equations		
(1)	$om(2)^* = -0.1043 + 0.01764d^2h$	0.985	19.4
(2)	$om(2) = -1.8161 + 0.01961d^2h - 0.0548d^2 + 0.3747h$	0.985	19.4
	Weighted least squares equations		
(3)	$om(2)/gh = 228.417 + 699.421(1/d^2h)$	0.847	12.6
	or		
	$om(2) = 0.0549 + 0.01794d^2h$	-	-
(4)	$om(2)/gh = 207.447 + 587.70(1/d^2h) + 208.05(1/h)$	0.876	11.4
	or		
	$om(2) = 0.0460 + 0.01629d^2h + 0.01634d^2$	-	-
(5)	$om(2)/gh = 218.943 + 311.286(1/d^2)$	0.898	10.3
	or		
	$om(2) = 0.017196d^2h + 0.024448h$	-	-
(6)	$om(2)/gh = 218.461 + 309.837(1/d^2) + 5.3084(1/h)$	0.898	10.3
	or		
	$om(2) = 0.01716d^2h + 0.02433h + 0.0004169d^2$	-	-

*om(2) = ovendry mass of stem wood + bark

Table 4. The fit of five regional equations for estimating the ovendry mass for six regional samples from across Canada

Sample location	dbh range (cm)	No. of sample trees	Equations* (1)	(2)	(3)	(4)	(5)**
			Estimated ovendry mass of stem wood plus bark as percent of observed values				
Alberta	0.1 - 10.0	102	100.1	96.7	94.4	93.5	91.0
	10.1 - 16.0	84	100.7	102.7	100.2	97.4	87.0
	16.1 - 32.0	66	100.8	105.7	103.1	99.8	84.6
	All	252	100.7	104.4	101.9	98.8	85.6
Ontario (1)	4.1 - 12.0	45	99.5	101.9	98.4	97.4	84.7
Great Lakes and	12.1 - 18.0	39	97.5	102.4	99.9	96.9	76.3
St. Lawrence Forest	18.1 - 44.0	40	95.3	101.2	98.8	95.5	76.6
Region	All	124	95.9	101.5	99.0	95.9	77.0
Ontario (2)	4.1 - 14.0	34	105.1	110.1	107.4	104.7	85.4
Boreal Forest	14.1 - 20.0	27	100.8	105.6	103.0	99.8	80.2
Region	20.1 - 36.0	32	95.3	100.6	98.1	94.9	80.2
	All	93	97.1	102.3	99.8	96.6	80.6
Ontario (3)	0.1 - 2.0	31	368.1	68.3	66.5	159.5	78.8
PNFI	2.1 - 4.0	62	146.4	92.0	89.8	111.6	96.8
	4.1 - 6.0	17	115.6	101.0	98.5	103.3	101.1
	All	110	147.9	94.3	92.0	112.1	97.4
Newfoundland	2.1 - 14.0	25	107.3	102.6	100.1	98.5	98.0
Central	14.1 - 22.0	21	97.3	100.5	98.1	95.0	85.4
	22.1 - 46.0	24	102.4	107.8	105.2	101.7	96.4
	All	70	101.4	105.9	103.3	99.9	93.8
Nova Scotia	0.1 - 12.0	16	95.9	85.2	83.1	82.6	92.3
Cumberland	12.1 - 22.0	14	99.3	100.9	98.3	95.3	94.5
County	22.1 - 34.0	16	106.0	110.5	107.8	104.3	105.9
	All	46	103.9	107.2	104.5	101.3	102.6

*(1) $om(2) = 1.6176 + 0.311d + 0.0178d^2 - 0.0005d^3 - 0.3916h + 0.0167d^2h$ (Alberta)
(2) $om(2) = 0.01794d^2h$ (Ontario (1))
(3) $om(2) = 0.01750d^2h$ (Ontario (2))
(4) $om(2) = 0.2420 + 0.01691d^2h$ (Newfoundland)
(5) $\ln om(2) = -3.0560 + 2.3536 \ln d + 0.2326 \ln h$ (Nova Scotia)
** Eq. (5) was used as given, without a correction factor (=1.01) as suggested by the authors, because its use would not have noticeably changed the estimates obtained.

Table 5. Regression data on the production of ovendry mass of trembling aspen, based on 675 observations from six regional samples from across Canada

	Equation	R^2	SEE %
Stem wood	$om(1)/gh = 181.987 + 181.911(1/d^2) - 39.5058(1/h)$	0.559	11.0
	or		
	$om(1) = 0.014293d^2h + 0.014287h - 0.0003103d^2$	-	-
Stem wood + bark	$om(2)/gh = 213.425 + 280.854(1/d^2) + 95.0715(1/h)$	0.735	11.1
	or		
	$om(2) = 0.01676d^2h + 0.022058h + 0.0074669d^2$	-	-
Stem wood + bark + live branches	$om(3)/gh = 225.672 + 246.244(1/d^2) + 303.008(1/h)$	0.704	12.4
	or		
	$om(3) = 0.017724d^2h + 0.01934h + 0.023798d^2$	-	-
Stem wood + bark + live branches + twigs & leaves	$om(4)/gh = 225.033 + 244.864(1/d^2) + 434.416(1/h)$	0.708	13.5
	or		
	$om(4) = 0.01767d^2h + 0.01923h + 0.034119d^2$	-	-

Table 6. The fit of four national equations for estimating the ovendry mass of trembling aspen for six regional samples from across Canada

Sample location	dbh range (cm)	No. of sample trees	Equations* (1)	(2)	(3)	(4)
			Estimated ovendry mass as percent of observed values			
Alberta	2.1 - 10.0	90	99.8	97.9	103.9	106.0
	10.1 - 16.0	82	102.7	100.3	105.2	107.0
	16.1 - 32.0	67	103.2	101.5	101.9	102.9
	All	239	102.9	101.0	102.7	104.0
Ontario (1)	4.1 - 12.0	45	99.0	100.3	101.8	98.0
Great Lakes -	12.1 - 18.0	39	99.7	98.9	102.6	101.5
St. Lawrence Forest	18.1 - 44.0	40	98.6	97.5	95.6	95.1
Region	All	124	98.9	97.9	97.2	96.4
Ontario (2)	4.1 - 14.0	34	101.7	107.4	114.0	113.4
Boreal Forest	14.1 - 20.0	27	98.6	101.8	106.3	106.1
Region	20.1 - 36.0	32	94.1	96.4	96.5	95.9
	All	93	95.6	98.3	99.7	99.2
Ontario (3)	0.1 - 2.0	30	100.9	101.9	105.4	104.9
PNFI	2.1 - 4.0	62	104.9	105.2	105.8	106.3
	4.1 - 6.0	17	105.8	105.8	104.1	103.9
	All	109	105.0	105.3	105.0	105.4
Newfoundland	2.1 - 14.0	23	98.7	101.5	106.9	102.6
Central	14.1 - 22.0	21	96.5	97.1	99.5	99.8
	22.1 - 46.0	24	102.7	103.5	101.6	101.3
	All	68	101.1	101.9	101.3	101.0
Nova Scotia	2.1 - 14.0	14	100.0	101.1	103.9	104.8
Cumberland County	14.1 - 22.0	12	104.5	100.5	99.2	100.5
	22.1 - 34.0	16	105.2	106.4	97.8	98.9
	All	42	104.8	104.9	98.3	99.4

*(1) $om(1) = (0.014293d^2h + 0.014287h - 0.0003103d^2) \times 1.002$
(2) $om(2) = (0.01676d^2h + 0.022058h + 0.0074669d^2) \times 1.003$
(3) $om(3) = (0.017724d^2h + 0.01934h + 0.023798d^2) \times 1.028$
(4) $om(4) = (0.01767d^2h + 0.01923h + 0.034119d^2) \times 1.028$

APPENDIX 2:

A note on the use of $s^2_{y \cdot x}$ and R^2 to compare equations

Appendix 2. A Note on the Use of $s^2_{y.x}$ and R^2 to Compare Equations

D.A. MacLeod
Computing and Applied Statistics Directorate
Place Vincent Massey, Hull, Quebec

When linear regression is employed to fit an equation of the form:

$$Y = a + b_1 X_1 + b_2 X_2 + ... + b_m X_m$$

to a set of data, the values of a, b_1, b_2 ..., b_m that are obtained are the best values in a specific sense; namely that RSS, the sum of squares of the residuals, are minimized. Other quantities obtained are the residual mean square $s^2_{y.x}$, given by

$$s^2_{y.x} = RSS/(n-m) \text{ where } n = \text{no. of data points}$$

and the multiple correlation coefficient R^2 given by:

$$R^2 = 1 - RSS/SS(Y)$$

where SS(Y) is the sum of squares of the deviations $Y_i - \bar{Y}$. Since RSS is always less than SS(Y) and represents the portion of SS(Y) that is not explained by the regression, R^2 can be considered the fraction of SS(Y) that is explained.

Considering R^2 from another angle, if the simple equation Y = a were fitted to the data, the value obtained for a would be $\bar{Y}$ and the deviations $Y_i - \bar{Y}$ would be the residuals. This means that R^2 represents the fraction of SS(Y) that is explained when the terms ($b_1 X_1 + b_2 X_2 + ... b_m X_m$) are added to the simple equation Y = a.

Let om, d and h represent the biomass, diameter and height of a tree respectively. For the three equations being considered in this report:

$$om = a + b(d^2 h) \qquad (1)$$

$$om = b(d^2h) \qquad (2)$$

$$om/d^2h = a + b/d^2 \qquad (3)$$

each equation makes different assumptions. One of these differences is the form of the equation - equ. (1) expresses om as a standard linear function of d^2h, equ. (2) is a straight line through the origin (i.e. om = 0 when $d^2h = 0$) and equ. (3), when each side is multiplied by d^2h, gives the more complex form $\underline{om = a(d^2h) + bh}$ which also goes through the origin. Another difference is the scale of the residuals -equations (1) and (2) assume that the residuals of om should be minimized while equ. (3) assumes that those of om/d^2h should be minimized.

If the equations are to be compared as to their goodness of fit to a set of data, using either $s^2_{y.x}$ or R^2, problems will arise from these differences in the assumptions. Equations (1) and (2) may be readily compared using $s^2_{y.x}$ (provided that the correct formula of RSS/(n-1) is used for equ. (2)), but neither can be compared to that for equ. (3) because the residuals are on a different scale. It is possible to transform equ. (3) and obtain residuals on the om scale, and from these get a transformed $s^2_{y.x}$. But this will put equ. (3) at a disadvantage since it was obtained by minimizing residuals on the scale of om/d^2h, not of om.

If the fit on one set of data is to be compared to the fit on another, a further difficulty will be present if, as is likely, the residuals tend to be larger for the larger trees. Here $s^2_{y.x}$ would represent an overall average and not the value at each point in the range of d or h. For the $s^2_{y.x}$ of one data set to be comparable to that of another, the data sets would have to occur in the same range of d and h.

Although there would be problems making comparisons using $s^2_{y.x}$, the problems using R^2 are much greater. In fact, no useful comparisons can be made among equations (1), (2) or (3) using the R^2 values output by standard computer programs.

The R^2 value for equation (1) is valid, representing the improvement when $b(d^2h)$ is added to om = a. But for equation (2) the value of a is set at zero. Most computer programs do not adjust for this, and simply calculate the improvement when $b(d^2h)$ is added to om = 0. Since the equation om = 0 does not fit the data at all (in contrast to om = a), there is a tremendous improvement when $b(d^2h)$ is added. The resulting R^2 is artificially large.

On the other hand, the R^2 for equation (3) is artificially small. It represents the improvement when b/d^2 is added to $om/d^2h = a$. Multiplying through by d^2h shows that this is the improvement when bh is added to $om = (d^2h)$. But the equation $om = a(d^2h)$ is already a reasonably good fit, so that the improvement when bh is added would not be that large. This would give a low R^2 even though the fit may be very good. A further complication is that this improvement is calculated on the scale of om/d^2h, not on the scale of om as it is for equations (1) and (2).

A possible solution, although it is not ideal, is to modify the R^2 to make it more comparable. It is first necessary to decide on what scale the comparisons will be made. For simplicity, assume that the om scale is selected. Then obtain a residual sum of squares RSS in that scale for each equation, transforming the equations where necessary, and define a modified R^2 to be

$$R^2(\text{mod}) = 1 - RSS/SS(om)$$

where SS(om) is the sum of squares of the deviations $om_i - \overline{om}$. What this does is to compare each equation (after transforming it if necessary so that om is alone on the left side) with the equation om = a. For equation (1), R^2(mod) is simply the ordinary R^2 value. For equations (2) and (3), extra calculation would be necessary to get R^2(mod).

Finally, it should be noted that R^2 (mod) compares equations at the level of the individual tree using the deviations $(om_i - \overline{om})$. In practice, however, the ultimate use of the equations is to obtain the biomass of stands rather than of individual trees. There is no guarantee that the equation that fits best on a tree basis will also give the best results on a stand basis. This can only be determined by applying the equations to stands of known biomass as is done in this report.