CRITICAL ANALYSIS OF MULTIVARIATE TECHNIQUES FOR ESTIMATING TREE TAPER SUGGESTS THAT SIMPLER METHODS ARE BEST

By A. KOZAK and J. H. G. SMITH¹

ABSTRACT

Trials of multivariate and other methods for analysis of tree taper are described. It is concluded that use of simple functions, sorting, and graphical methods is adequate for many uses in operations and research.

Fries (1965) and Fries and Matern (1965) introduced multivariate methods for construction of an integrated system of models of tree taper curves. They expected that expression of these models in mathematical functions would have considerable advantage over those given only as tables or graphs. Grosenbaugh (1966) observed that polynomial analysis may rationalize observed variation in form after measurement, but does not promise more efficient estimation procedures.

The method of Fries (1965) was applied to determine eigenvalues from representative samples of B.C. Forest Service data for all major groups of the commercial tree species of British Columbia. The basic data were chosen from trees measured in volume and decay studies to cover a wide range of tree size, site quality, and locality. Trees had been fully described and measured after felling to find diameters inside and outside bark (dib and dob) at stump height, breast height, and by sections cut at each tenth of tree height above 4.5 feet. The tree sizes represented are shown by species groups in Table 1. Percentage distribution of trees sampled is given by taper classes in Table 2.

The method used by Fries (1965) calculates elements of the first vector which are analogous to a freehand line drawn through measurements of dib plotted over the percentages of total tree height corresponding to stump, bh, and other sectional values. The second vector is analogous to a freehand line drawn through the residuals of dib from the first line and plotted over the same percentages of total height. The process of calculating eigenvectors is repeated until all residuals are accounted for, or the number of vectors equals the number of degrees of freedom. The eigenvalues and elements for each fraction of total height are all relative and are determined with appropriate weighting from the data by application of matrix algebra. Constants for each tree are calculated and these are summarized by convenient groupings related to determinants of tree form. Constants required for definition of taper of each tree can be estimated by multiple regression techniques based on individual tree data for dbh (D), total height (H), and H/D, once the basic data have been analyzed.

¹ Respectively, Asst. Professor, and Professor, Faculty of Forestry, Univ. of B.C., Vancouver, B.C. Financial support of the National Research Council of Canada and provision of data by the B.C. Forest Service are gratefully acknowledged.

About 99.7 per cent of the variation in dib (eigenvalues) is accounted for by the elements of the first eigenvector. This percentage ranged from 99.3 to 99.9 for the species groups in Table 1. The second eigenvector accounted for about 0.15 per cent, and vectors three to ten for the remaining 0.15 per cent.

Most of the variation in the elements of eigenvector No. one is associated with tree size. Data for the median class have been summarized in Table 3. The differences in elements of the matrix are greatest absolutely and relatively near the tip. However, the variations in tree shape are relatively minor and not strongly controlled by species for these species groups. Even within the extremes of taper represented by the 500 D/H classes 0 to 65 and 75 to 100 which are shown in table 4 the differences are not great. The averages shown in table 4 indicate the variation that can be estimated by tree variables such as D and H.

TABLE 1
SUMMARY OF TREE DATA USED

Species group ¹	No. Trees	Height Ft.			Dbh In.			H/D Ft./In.
		Min.	Ave.	Max.	Min.	Ave.	Max.	
C. Douglas fir	97	52	122	231	6	22.4	68	7.50
I. Douglas fir	78	27	75.9	120	5	17.3	38	5.82
C. Spruce	39	48	143	246	6	29.5	85	5.68
I. Spruce	121	37	89.3	137	5	13.3	26	7.05
C. Balsam	86	32	104	184	6	18.0	50	6.31
I. Balsam	96	28	69.9	118	5	10.3	24	6.98
C. Red cedar	85	31	89.6	189	7	26.2	77	5.43
I. Red cedar	91	27	80.2	162	6	21.6	62	6.08
C. Hemlock	81	39	99.6	176	5	20.3	66	7.45
I. Hemlock	63	42	72.9	125	6	12.9	38	7.02
Yellow cedar	74	38	65.8	148	6	16.9	54	5.25
I. White pine	95	45	107	147	6	19.2	34	7.30
I. Yellow pine	101	20	78.1	135	6	19.2	42	5.34
I. Lodgepole pine	456	42	75.3	130	5	10.7	23	7.31
Black cottonwd	142	45	80.6	167	5	12.7	82	8.04
Alder	101	53	83.2	108	5	12.7	30	7.91
Aspen	126	40	77.7	103	5	12.4	29	7.88
Maple	52	48	77.5	95	6	14.1	30	7.07

¹ C is Coastal, I is Interior, British Columbia.

TABLE 2
DISTRIBUTION OF TREES SAMPLED BY CLASSES OF 500 D/H

Species group ¹	No. Trees	500 dbh/Height Percentage in class				
		0-65	65-8	68-71	71-75	75-100
C. Douglas fir	97	16	9	13	18	42
I. Douglas fir	78	14	8	12	17	49
C. Spruce	39	82	5	13	—	—
I. Spruce	121	35	15	19	16	15
C. Balsam	86	35	15	19	16	15
I. Balsam	96	30	19	18	15	18
C. Red cedar	85	69	12	7	5	7
I. Red cedar	91	67	9	8	8	8
C. Hemlock	81	46	16	10	14	14
I. Hemlock	63	25	11	18	23	23
Yellow cedar	74	24	22	19	16	19
I. White pine	95	38	26	20	9	7
I. Yellow pine	101	5	10	11	13	64
I. Lodgepole pine	456	27	14	16	13	30
Black cottonwood	142	47	22	18	11	2
Alder	101	17	8	23	37	15
Aspen	126	18	20	26	23	13
Maple	52	57	12	4	10	17

¹ C is Coastal, I is Interior, British Columbia.

TABLE 3
INFLUENCE OF SPECIES ON ELEMENTS OF EIGENVECTOR No. 1.

Percentage of total height of tree above ground	Elements of dib matrix for classes of 500/H: 65 to 68 all species tested			
	Min.	Max. Element for each section	Diff.	Ave.
4	474 ¹	516	42	488
10	412	431	19	423
20	419	445	26	429
30	387	413	26	399
40	358	374	16	366
50	322	332	10	328
60	269	291	22	282
70	209	243	34	227
80	144	181	37	163
90	73	146 ¹	73	90

¹ Maple was very much different from the other species.

TABLE 4
INFLUENCE OF TAPER ON ELEMENTS OF EIGENVECTOR NO. 1

Percentage of total height of tree above ground	Elements of dib matrix for classes of 500 D/H: Differences within species for the classes 0 to 65 minus 75 to 100 for all species tested		
	Min.	Max.	Ave.
4	26	60	39
10	32	98	52
20	8	28	16
30	— 9	13	3
40	— 1	—27	—13
50	—21	—58	—32
60	—21	—45	—32
70	—22	—48	—34
80	— 8	—51	—27
90	0	—52	—18

The elements of the vectors corresponding to eigenvalues two and three were much more variable than those for eigenvalue one. The very small fractions of variation accounted for by them was seldom related in a consistent manner to the tree variables available in the B.C. Forest Service data. Estimations of the functional form of the eigenvectors AK1, AK2, and AK3 (Fries and Matern, 1965) to define eigenvectors one, two, and three were made for 456 lodgepole pines. Regression estimates were made for tree height, age, dbh, and their combinations D/H, D/A, H/A, D²H, and H/D. Various trials of these for estimation of dib by classes of D and H showed that the improvement gained from use of eigenvectors two and three was slight and not justified by the extra work.

In an attempt to refine the estimation of AK1, AK2, and AK3, 102 of the lodgepole pine analysed by Stanek (1966) were studied. Multiple regression analysis of 22 variables for each tree showed that dbh accounted for 98.8 per cent of the variation in AK1. Addition in turn of Hohenadl's form constant Lambda 0.9 (Heger, 1965), bark percentage, and H/D increased the variation accounted for by only 0.4 per cent. In estimation of AK2, Lambda 0.9 accounted for 57.2 per cent of the variation which was increased to 70 per cent by addition of D/A and H/A. Estimation of AK3 was much poorer. Lambda 0.9 accounted for 5.7 per cent, which could be increased to 12.9 per cent by addition of crown width. All 22 variables accounted for 99.4 per cent of the variation in AK1, 80.7 per cent in AK2, and 25.6 per cent in AK3. Again the use of many variables and inclusion of AK2 and AK3 did not seem justified by the minor improvement in estimation of dib for each section.

Conventional regression methods (Kozak and Smith, 1965) and a large, principal components regression program with varimax rotation of the factor weight matrix (Wallis, 1965) were also used to estimate dib at various heights directly for our original 456 lodgepole pine trees. Results obtained by both approaches were quite comparable with the same regression coefficients being

calculated to at least the fifth decimal place. For application of the principal components program, dib at each section height was estimated in terms of D , H , and H/D . The regression constants obtained for each section then were plotted and an appropriate function was calculated to describe the variation of these constants with position in tree. The mathematically harmonized constants were used to estimate the dib of each of the sections for each of 456 trees. Comparison of actual with estimated dib for each of the 4,560 sections showed a standard error of 1.063 inches.

Since inspection of graphs of diameter-squared plotted over section height usually confirms the adequacy of the assumption that the quadratic parabola applies well, several simple approaches were also tried. Honer (1964) described some easy methods for estimation of merchantable cubic foot volume. Smith and Munro (1965) tested these and other approaches to define distribution of dib and volume within trees. In subsequent analyses they showed, for the commercial trees of B.C., that section volumes up to heights h , expressed as fractions of volume per tree could be estimated with a standard error of estimate (S.E.E.) of 10 per cent by use of h/H and with an S.E.E. of 3 per cent by use of h/H and h^2/H^2 . Munro (1966) developed a new linear transformation of the quadratic paraboloid which describes very well the distribution of dib throughout most of our trees. Although it fails to describe tree shape fully at the butt and tip it satisfies many of our needs for estimation of fraction of volume merchantable to various specifications of dib and h .

Application of Munro's equation to estimate dib of each section of our 456 lodgepole pine trees showed a standard error of only 0.662 inches. The equation used was:

$$d = D \sqrt{1.07923 - 1.1175 \frac{h}{H - 4.5}}$$

where d is dib, D is dbh(ob), h is height to the section, and H is total height.

Deviations from Munro's equation might be explained partly by further analysis of various measures of crown development but since much of the variation seems to be associated with random variations in knot clustering and bole cross-sectional shape little improvement can be expected.

Our trials have shown that for operational purposes the taper curves prepared by the B.C. Forest Service for all of its commercial species groups can be used with considerable confidence. If taper tables are required they can be prepared from our data easily by appropriate sorting techniques on an electronic computer. Since our analyses have shown the similarity of tree shape of the commercial tree species of B.C., and the difficulty of refining taper curves further by multivariate analyses for lodgepole pine and other species, we conclude that use of simpler methods is best. However, we may continue to study the problem with many new ideas gained from Grosenbaugh's (1966) comprehensive review of tree form.

REFERENCES

- FRIES, J. 1965. Eigenvector analyses show that birch and pine have similar form in Sweden and in British Columbia. *For. Chron.* 41(1) 135-9.
 FRIES, J. and B. MATÉRN, 1965. On the use of multivariate methods for the construction of tree taper curves. I.U.F.R.O. Advisory Group of Forest Statisticians Conference, Stockholm, 33 p. Mimeo.

- GROSENBAUGH, L. R. 1966. Tree form: Definition, interpolation, extrapolation. *For. Chron.* 42(4): 443-56.
- HEGER, L. 1965. A trial of Hohenadl's method of stem form and stem volume estimation. *For. Chron.* 41(4): 466-75.
- HONER, T. G. 1964. The use of height and squared diameter ratios for the estimation of merchantable cubic foot volume. *For. Chron.* 40(3): 324-31.
- KOZAK, A. and J. H. G. SMITH. 1965. A comprehensive and flexible multiple regression program for electronic computing. *For. Chron.* 41(4): 438-43.
- MUNRO, D. D. 1966. The distribution of log size and volume within trees: A preliminary investigation. Univ. of B.C., Fac. of For. Directed Study 27 p.
- SMITH, J. H. G. and D. D. MUNRO. 1965. Point sampling and merchantable volume factors for the commercial trees of British Columbia. Univ. of B.C., Fac. of For., 39 p. Mimeo.
- STANEK, W. 1966. Occurrence, growth, and relative value of lodgepole pine and Engelmann spruce in the Interior of British Columbia. Univ. of B.C., Fac. of For. Ph.D. thesis, 252 p. Mimeo.
- WALLIS, J. R. 1965. WALLY1 . . . A large, principal components regression program with varimax rotation of the factor weight matrix. U. S. F. S. Res. Note PSW-92. 6 p.