MATCHACURVE — 4:

Segmented Mathematical Descriptors for Asymmetric Curve Forms

Chester E. Jensen

USDA Forest Service Research Paper INT - 182
INTERMOUNTAIN FOREST AND RANGE EXPERIMENT STATION

USDA Forest Service
Research Paper INT-182
1976

MATCHACURVE — 4:
Segmented Mathematical Descriptors for Asymmetric Curve Forms

Chester E. Jensen

INTERMOUNTAIN FOREST AND RANGE EXPERIMENT STATION
Forest Service
U.S. Department of Agriculture
Ogden, Utah 84401

THE AUTHOR

CHESTER E. JENSEN is Principal Statistician for the Intermountain Forest and Range Experiment Station. Previously, he held the same position at the Northeastern and Central States Forest Experiment Stations. Mr. Jensen, who holds a master of forestry degree from Michigan State University, received graduate training in statistics at Iowa State University.

CONTENTS

ABSTRACT

Methods are shown for achieving sensitivity in mathematically describing asymmetric curve forms. Independent, but compatible, descriptors are developed for each of several response surface regions or segments, in the examples given. The methods are analogous for problems necessitating the use of additional segments and/or dimensions.

INTRODUCTION

Mathematical description of asymmetric curve forms expected in a regression relation can be difficult when descriptor components must apply over the entire range of each independent variable involved. Under such constraint, failure of the analyst to find suitably accurate forms (Bartlett [1947] and Draper and Hunter [1969] recognized the necessity for allowing the analyst to establish his own acceptance criteria) among his available alternatives might prompt adoption of a segmented descriptor system. For our purposes, a segmented descriptor system is one wherein the relation to be described is divided into two or more segments covering the ranges of one or more independent variables. A descriptor is developed for the portion of the response curve within each segment and applies there exclusively. The contiguous array of these segments then portrays the entire relation. The segmented approach might be regarded as less elegant, perhaps, from the standpoint of mathematical and statistical manipulability, but this may be a necessary trade-off for descriptor accuracy.

Segmented descriptors are developed for several three-dimensional relations using systems outlined in Matchacurves-1 and -2 (Jensen and Homeyer 1970, 1971) and in Matchacurve-3 (Jensen 1973). The first example involves a two-segment descriptor. At the time of analysis, the general form of the relationship was known. A bell-shaped curve, possibly asymmetric, was expected over the range of one independent variable, while a sigmoidal curve was expected over the second. Strong interaction was likely to occur between the independent variables.

Specification of a viable mathematical hypothesis for this potentially complex form, from prior knowledge alone, was considered to be impractical. And, rather than dedicate the data to the statistical evaluation of poorly specified hypotheses and suffer a potential loss of information, the analyst elected to exhaust the data graphically. The graphed model was then described mathematically. The resulting function was refitted to response values at 36 control points on the graph, this by least squares in the simple model....

$$Y = \beta \text{ (model transform) } + \varepsilon$$

Example #2, a three-segment descriptor, was completely specified in graphic form from prior knowledge. The objective here was simply to describe the graphed model mathematically, using the Matchacurve system. Data were not involved.

Both the graphic and general mathematical forms of these models are presented in the text. Familiarity with the Matchacurve system will enhance understanding of the segmentation adopted. Specific mathematical forms and associated explanatory material are given in the appendix to reinforce the reader's knowledge, as needed, of mathematical component development by means of Matchacurve.

EXAMPLE 1
AN ASYMMETRICAL, BELL-SHAPED RIDGE: TWO SEGMENTS

In this descriptor, tree mortality percent in a western forest is characterized by tree diameter (d.b.h.) over the course of a beetle epidemic (fig. 1). The descriptor is applicable only at discrete points in time and at the midpoints of 2-inch d.b.h. classes. Both variables are treated as continuous. As was evident in the original data, the more-or-less bell-shaped trends over time differ substantially on either side of the central ridge. For example, observe strong asymmetry at d.b.h. = 12 to 18 inches, made most noticeable perhaps by elevational differences in the curves at their left and right extremes. See also the dwindling breadth of curve crowns above and below 14 inches d.b.h., an interactive change included in the descriptor along with asymmetry. Descriptor components were assembled as follows:

Left and right sides of the ridge were described separately (fig. 2 and 3), using sigmoids from Matchacurve-1 that varied in shape and scale according to data trends over d.b.h. Because the lower portion of the left half asymptoted at values larger than zero, it was necessary to include the left-edge intercept, Int, (or floor) upon which the left-side sigmoids rested conceptually. Int is a function of d.b.h. as is the location of the ridge in time, XP.

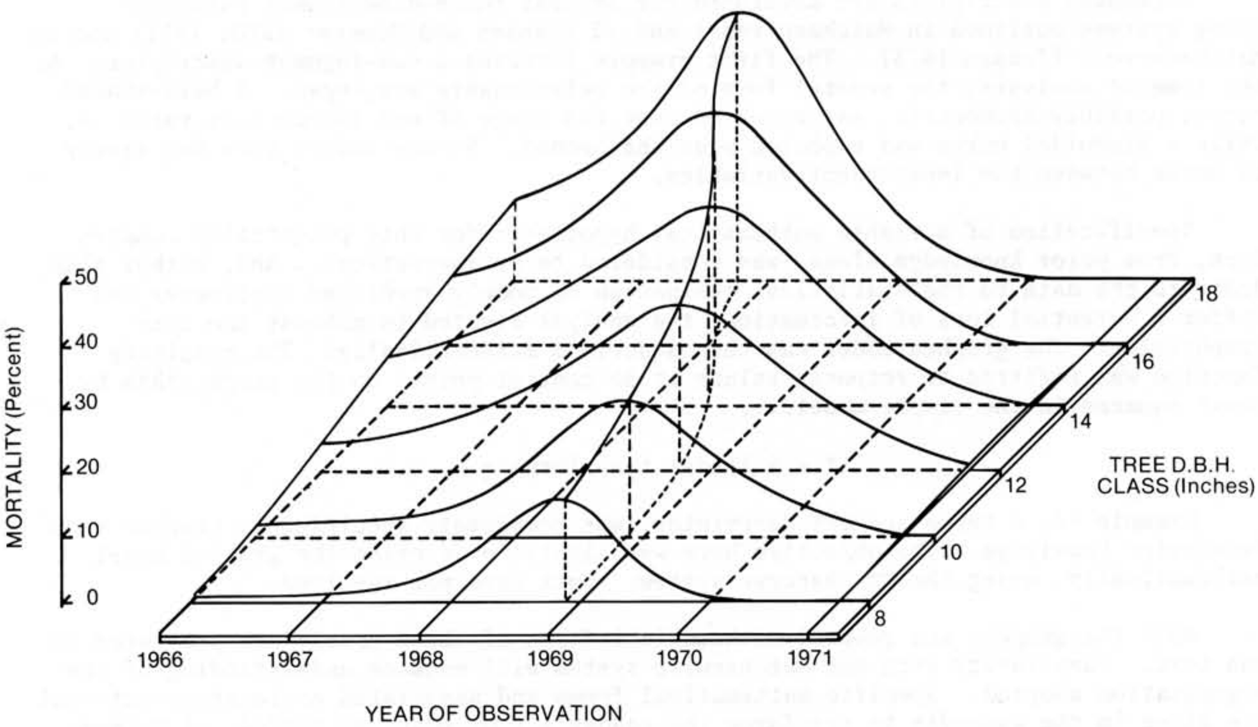

Figure 1.--Tree mortality over the course of a beetle epidemic.

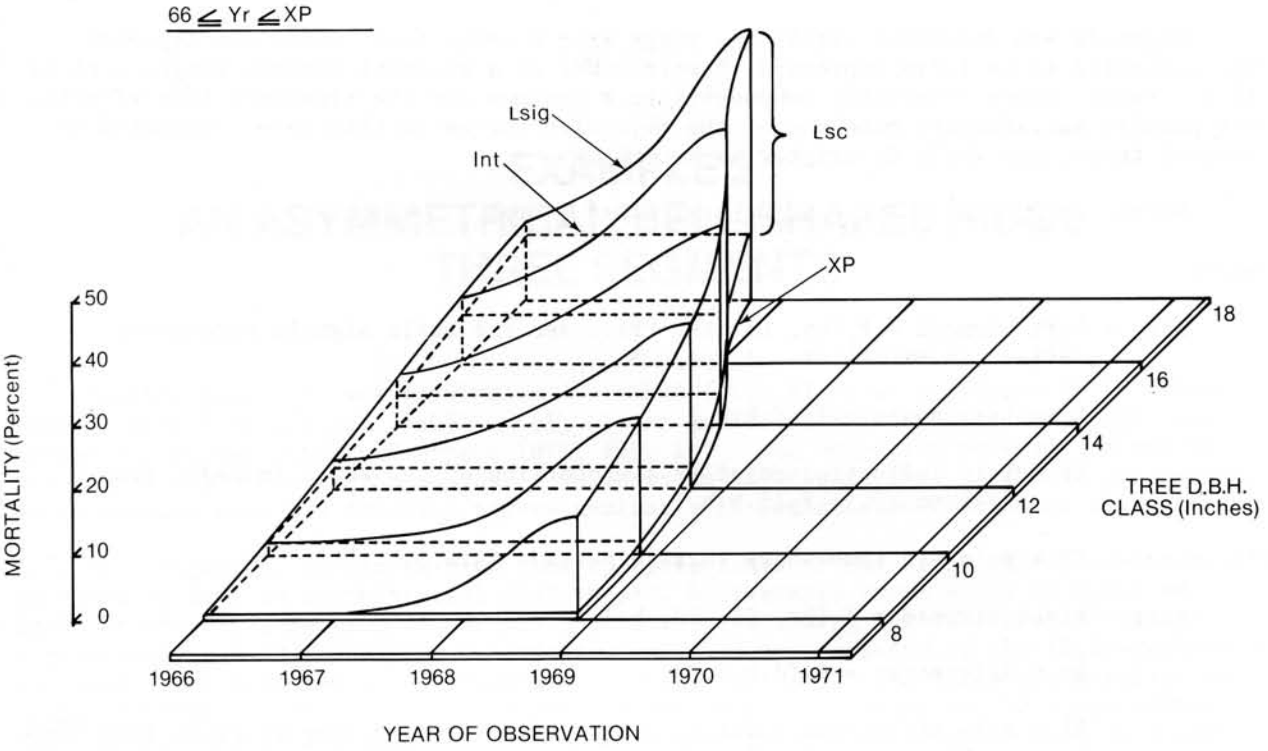

Figure 2.--Tree mortality over the course of a beetle epidemic....<u>left segment</u>.

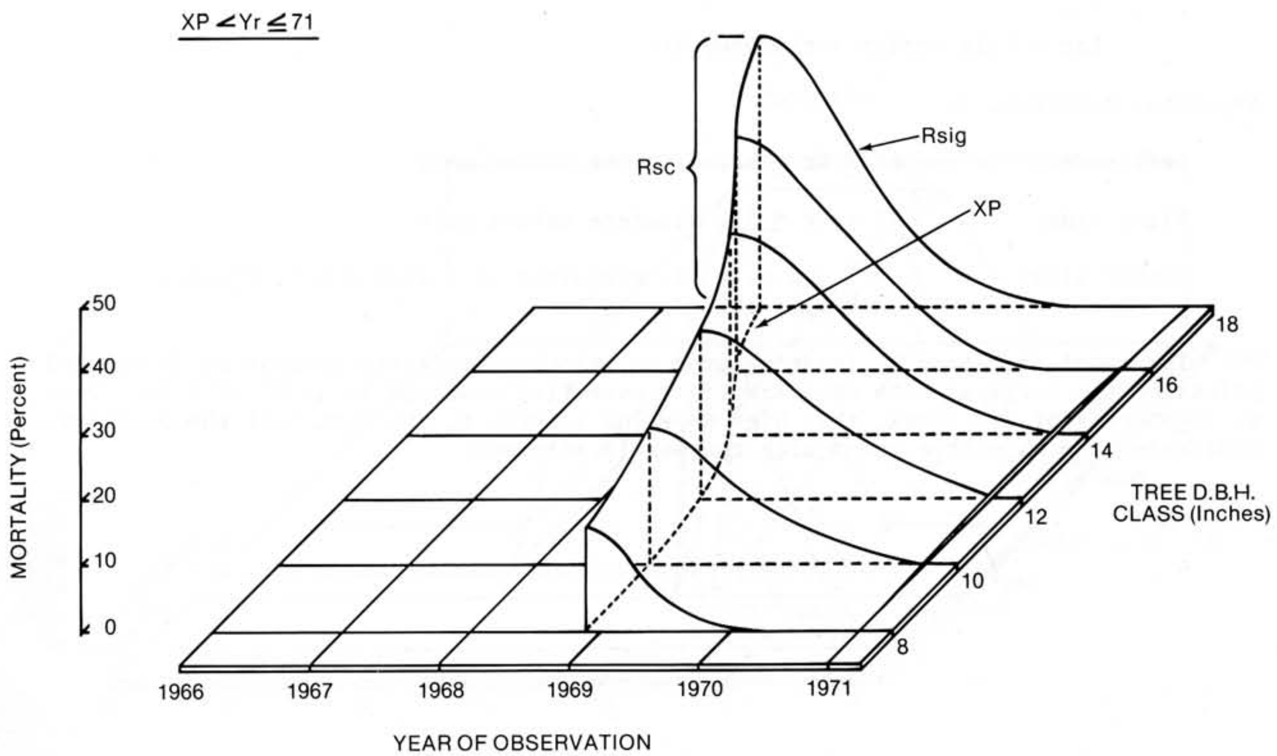

Figure 3.--Tree mortality over the course of a beetle epidemic....<u>right segment</u>.

Sigmoids are described within the range XP ± 5 years (yr), since all sigmoids are estimated to be fully expressed therein. Use of a constant maximum range, such as XP ± 5 years, keeps descriptor components to a minimum for the sigmoidal time effects, yet permits satisfactory matching of the objective curves in this case. Restated in general terms, the whole descriptor is as follows:

Mortality percent = Int + Lsc (Lsig) + Rsc (Rsig)

where

Lsig = left sigmoid = f_L(Ln, LI, XP, Yr)...see the basic sigmoid parameters defined in Matchacurve-1, page 3.

Ln = Lsig power = f_1(d.b.h.)

LI = Lsig inflection point as a proportion of the range in years from XP-5 to XP, = f_2(d.b.h.)

XP = point in time where surface peaks = f_3(d.b.h.)

Rsig = right sigmoid = f_R(Rn, RI, XP, Yr)

Rn = Rsig power = f_4(d.b.h.)

RI = Rsig inflection point as a proportion of the range in years from XP+5 to XP, = f_5(d.b.h.)

Int = intercept, left edge = f_6(d.b.h.)

Rsc = Rsig scalar = ridgetop = f_7(d.b.h.)

Lsc = Lsig scalar = ridgetop-Int

Segmental constraints:

Left side; $66 \le Yr \le XP$, discrete values only

Right side; $XP < Yr \le 71$, discrete values only

Either side; $8 \le d.b.h. \le 18$, midpoints of 2-inch d.b.h. classes only...8, 10, 12, etc.

The model, refitted by least squares to smoothed mortality percent at 36 control points on the original data cross-sections over time resulted in an R^2 of 0.96. Used as a goodness-of-fit index, this high R^2 value attests to the fact that the descriptor duplicates the objective graph with reasonable accuracy.

EXAMPLE 2
AN ASYMMETRICAL, BELL-SHAPED RIDGE: THREE SEGMENTS

Quality score (t) was originally described (fig. 4) as an aggregate of 11 planar regions over flow (F) and stability (S) of gap-graded road materials, a strongly segmented descriptor (Lee, and others 1973, fig. 14). The objective here was to *smooth* the figure and *minimize* the number of segments in the descriptor. Although the problem is an unusual one, its solution serves admirably to demonstrate descriptor segmentation.

From figure 4, it can be seen that opposite sides of the ridge differ substantially in slope so that an asymmetrical, flat-topped, bell-shaped curve would be required to describe the cross-section at any point in S while rounding the junctures of planes. A relatively simple descriptor alternative involves segmentation of the (S,F)-regions as shown, with left and right orientation lines. Note that the lines are parallel to their respective sides of the ridge and lie one unit in F closer to the ridge center. This permits use of a single, but different, sigmoid cross section to represent and smooth the corners of each segment, left and right. The sigmoids will peak at their respective orientation lines and will be functional over a constant distance from them (left or right as appropriate). The constant will vary by side as needed. Since the ridge is flat-topped, it will have a value of t = 10 everywhere in the center segment.

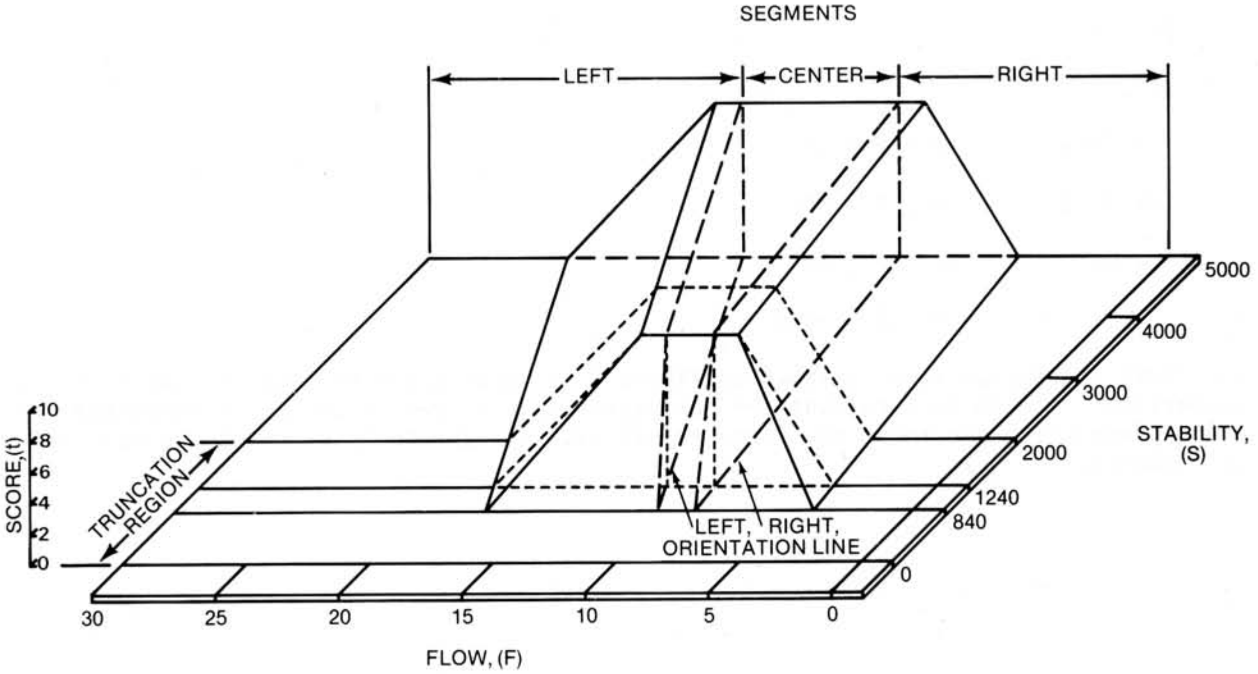

Figure 4.--Quality of gap-graded road materials: segmented planar form.

Then, to this point, we have described a ridge with three segments specified in terms of S and F: the sigmoidal effect to the left of the left orientation line; the sigmoidal effect to the right of the right orientation line; and the flat ridge area between these lines at a value of t = 10.

Finally, a sigmoidal truncation of the front end of this ridge is achieved through multiplication of all components of the descriptor by an appropriate sigmoid, changing in value from zero to one within the range $0 \leq S \leq 2000$, and being applicable for $0 \leq S \leq 5000$.

In general terms, we have:

Lsig and Rsig = left- and right-segment sigmoids, respectively

LO, RO = left- and right-orientation lines, respectively

CC = center segment, constant

Tsig = truncator sigmoid

Then

t = Tsig (Lsig + CC + Rsig)

and

Tsig = f(S)

Lsig = f(LO,F)

Rsig = f(RO,F)

LO = f(S)

RO = f(S)

Limits

For Lsig,	$LO < F \leq 28$
For Rsig,	$0 \leq F < RO$
For CC,	$LO \leq F \leq RO$
and,	$0 \leq S \leq 5000$

The final descriptor form is shown in figure 5 along with the original planar form for comparison. It can be seen that the descriptor does a creditable job of emulating the planar form while smoothing the corners, all with 3 segments in place of the original 11 segments.

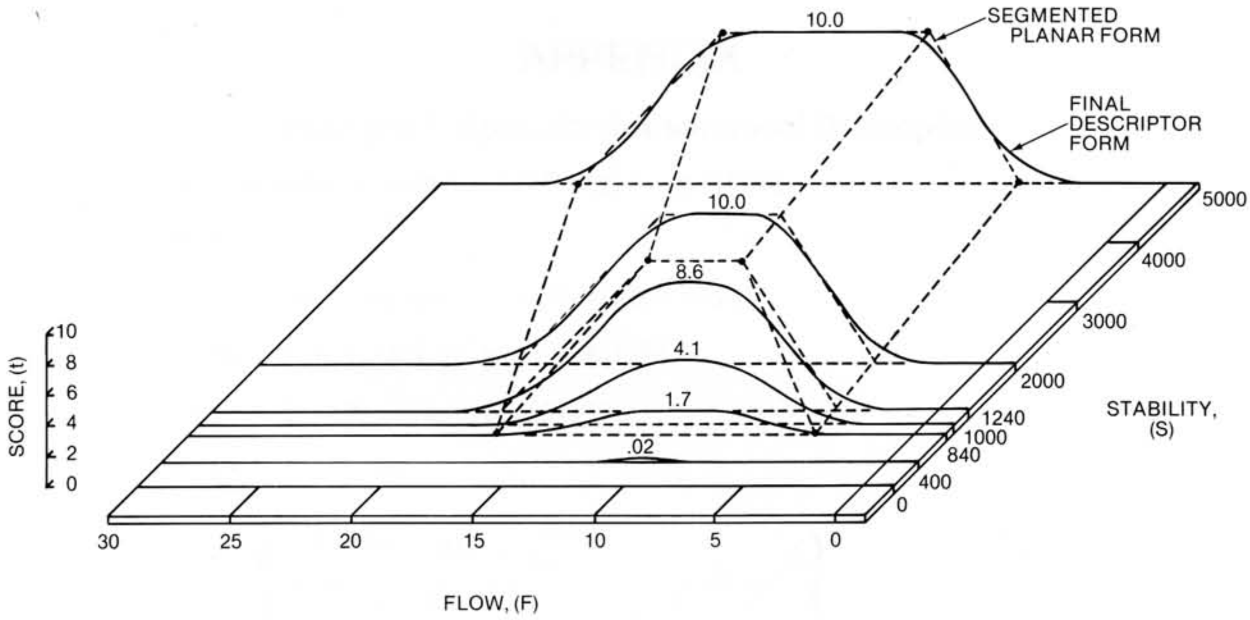

Figure 5.--Quality of gap-graded road materials: segmented planar and final descriptor forms.

APPENDIX

Example 1: Specific Mathematical Descriptor

Mort. percent = K(Int + Lsc[Lsig] + Rsc[Rsig])

where

K = least squares coefficient = 0.9877

$$\text{Int} = 1 + 0.2321\,(\text{d.b.h.} - 8)^{1.63}$$

$$\text{Rsc} = 2.6429\,(\text{d.b.h.}) - 5.157$$

Lsc = Rsc - Int

$$\text{Lsig} = \left\{ \frac{e^{-\left|\frac{\frac{5 - |\text{Yr} - \text{XP}|}{5} - 1}{1 - \text{LI}}\right|^{\text{Ln}}} - e^{-\left(\frac{1}{1 - \text{LI}}\right)^{\text{Ln}}}}{1 - e^{-\left(\frac{1}{1 - \text{LI}}\right)^{\text{Ln}}}} \right\}$$

$$\text{XP} = 67.65 + 1.2257\, e^{-\left|\frac{\frac{20 - \text{d.b.h.}}{12} - 1}{0.65}\right|^{3.9}}$$

$$\text{LI} = 0.851 - 0.1531\, e^{-\left|\frac{\frac{7 - |\text{d.b.h.} - 13.5|}{7} - 1}{0.45}\right|^{2}}$$

Ln = 1.5

Rsig = as for Lsig using RI and Rn

$$\text{RI} = 0.752 + 7.8173 \times 10^{-4}\,|\text{d.b.h.} - 14|^{2.8}$$

Rn = 2

Limits

$66 \le \text{Yr} \le 71$, discrete units only

$66 \le \text{Yr} \le \text{XP}$, for Int, Lsc, Lsig

$\text{XP} < \text{Yr} \le 71$, for Rsc, Rsig

$8 \le \text{d.b.h.} \le 18$

SUPPLEMENTARY EXPLANATORY DETAIL FOR EXAMPLE 1

Int was estimated from leftward extension, from the ridge, of the mortality trend indicated by data in each of six 2-inch d.b.h. groups. All six trends reached lower asymptotes by 1963 and at that point suggested a flat, concave-upward curve over d.b.h. (fig. 6). This was satisfactorily described using Matchacurve-2, set A-1.

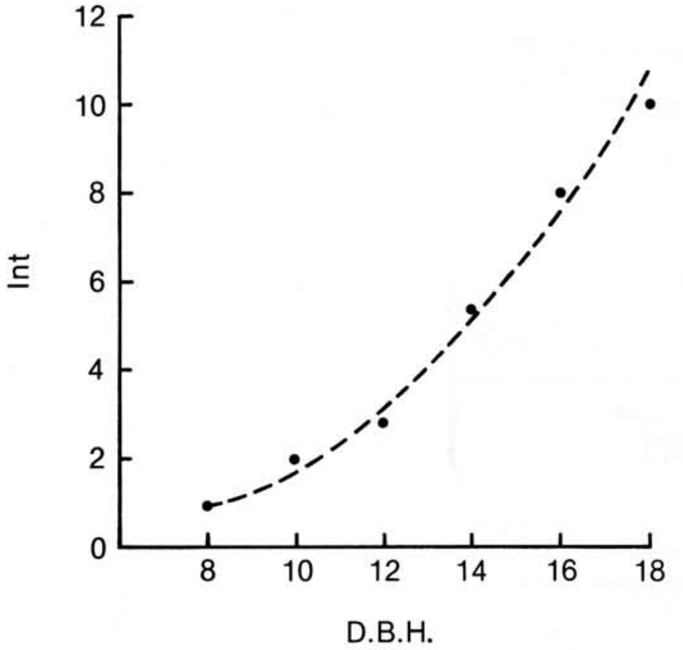

Figure 6

Rsc is simply the height of the ridge above zero and serves as the scalar for the right-half sigmoids. Rsc is a linear function of d.b.h., adopted to represent the somewhat irregular pattern of ridge values for the six d.b.h. groups (fig. 7). Note that the ridge line in figure 1 only *appears* to be sigmoidal by reason of the sigmoidal change of point-of-peaking in time, with a change in diameter.

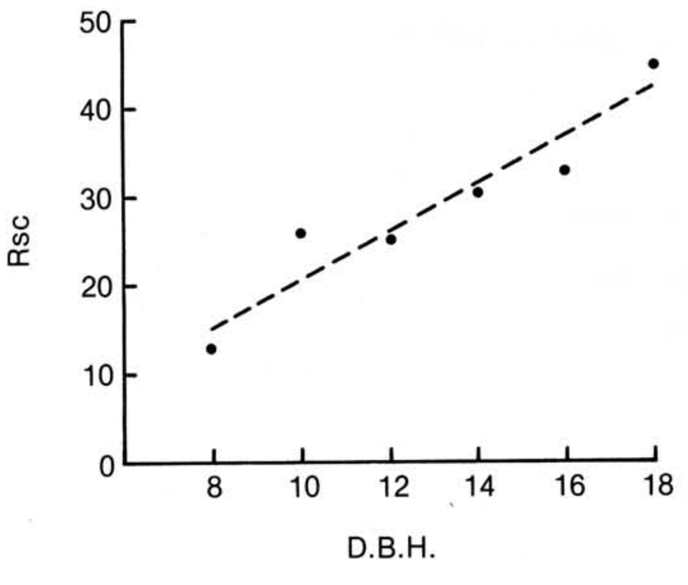

Figure 7

Lsc = (Rsc-Int), and is the scalar for the left-half sigmoids.

XP, the point in time at which the bell-shaped curves peaked, was estimated from the XP's for the six data-group cross-sections (fig. 8). This curve was estimated to asymptote at 68.88 and 67.75. A suitable match was found in the sigmoids of Matcha-curve-1 when the d.b.h. was reversed; namely, transformed to 20-(d.b.h.).

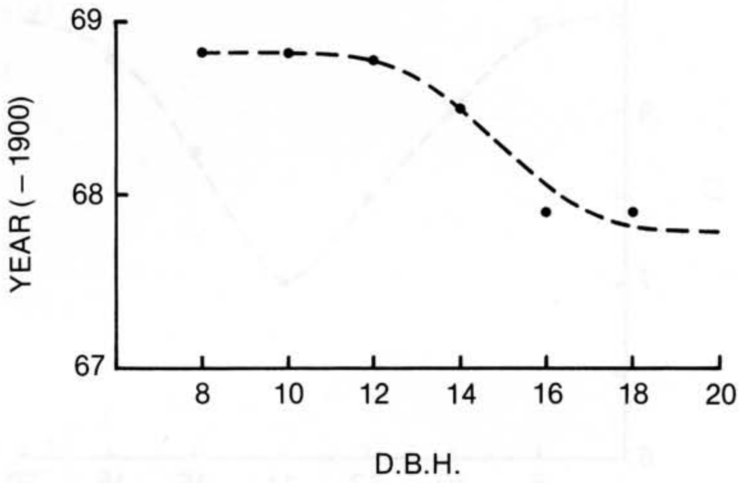

Figure 8

Lsig specifies the sigmoidal shape of the left side. After transforming Year to 5-|YR-XP|, to create X-values ranging from zero at XP-5, to 5 at XP, the six d.b.h.-group cross-sections (left halves) were each scaled to 1.0 in X and Y...at critical points in X. Overlay curves (see Matchacurve-1) were plotted for these cross-sections (fig. 9). The pattern was one of wider curve crowns near d.b.h. = 14 inches, narrowing with increasing departure from that d.b.h. This is reflected in the bell-shaped function for the inflection points, LI. Also, Matchacurve-1 Standards with n = 1.5 were found to represent this curve array with reasonable accuracy, so Ln was set at the constant, 1.5.

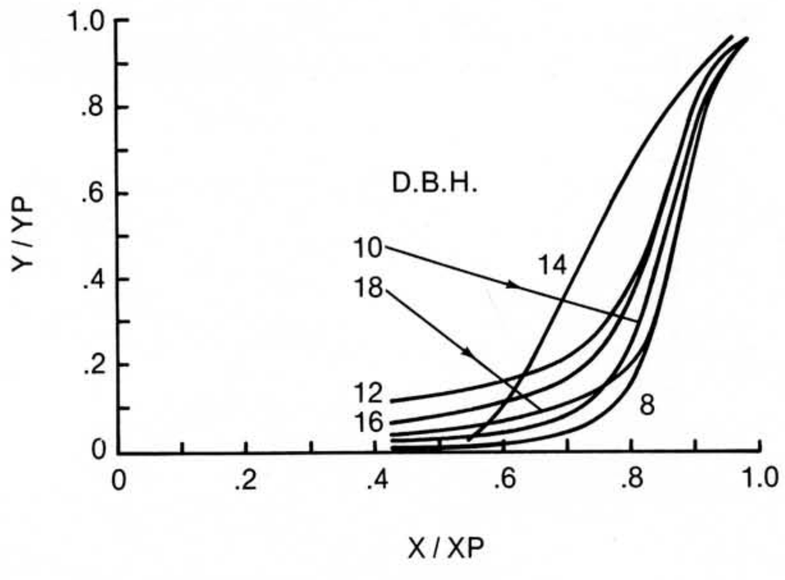

Figure 9

LI, the function for the inflection points, is the dashed line in figure 10, and represents the inflection points adopted and plotted for each overlay curve. LI reflects the width-of-curve-crown trend noted under Lsig and in the text.

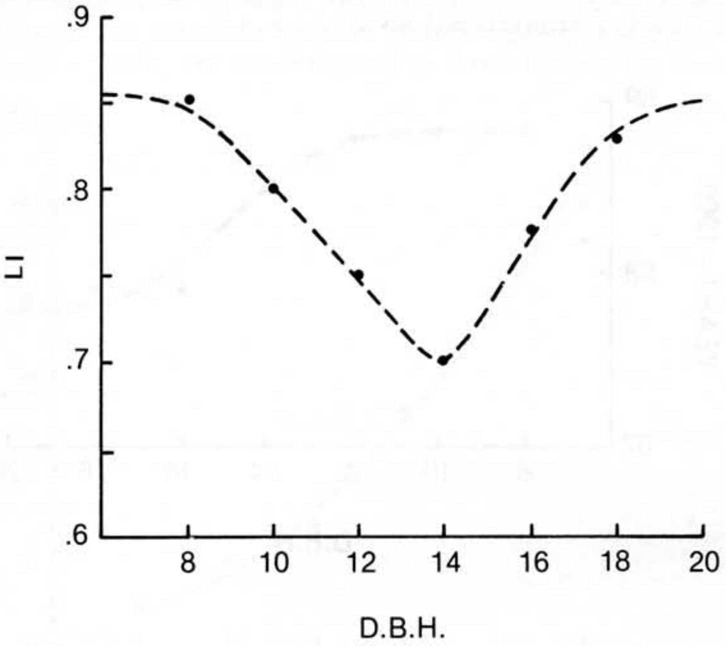

Figure 10

Rsig, RI, and Rn were obtained in a similar fashion. The Year transform again was 5-|Yr-XP| and ranged in value from 5.0 at XP to zero at XP + 5. The right-side cross-sections were fairly well matched by Standards with n = 2, so Rn = 2. RI showed the same general widening of curve crown at d.b.h. = 14 inches.

Example 2: Specific Mathematical Descriptor

If

$F > LO$, $t = Tsig(Lsig)$

$LO \le F \le RO$, $t = 10(Tsig)$

$F < RO$, $t = Tsig(Rsig)$

where

F = soil flow

$LO = 8.888 + 0.0017036(S)$

$RO = 8.404 + 0.00048076(S)$

S = soil stability

$$Lsig = 10\,e^{-\left|\frac{\frac{(LO + 20 - F)}{20} - 1}{0.23}\right|^{2.2}}$$

$$Tsig = 1 - e^{-\left|\frac{\frac{(6000 - S)}{6000} - 1}{0.185}\right|^{6}}$$

$$Rsig = 10\,e^{\left|\frac{\frac{|RO - 11| + F}{11} - 1}{0.30}\right|^{2.4}}$$

Limits

$0 \le F \le 28$

$0 \le S \le 5000$

SUPPLEMENTARY EXPLANATORY DETAIL FOR EXAMPLE 2

Lsig

The left segment of figure 4 has a constant cross-section over F for 1240 ≤ S ≤ 5000, approximated in the descriptor by a sigmoid oriented at F = LO and extending leftward a distance of 20 units in F. Twenty is about the minimum operational span for the sigmoids here. At the upper extreme of S, (S = 5000), the left-segment sigmoid is expected to be completely specified in the range 27 ≥ F ≥ LO, as in figure 5. Setting 28 as the upper level of F within which all left-segment sigmoids must be completely specified, we turn to the limiting case at S = 0. Here LO = 8.888 and the sigmoidal range must then be at least (28 - 8.888) = 19.112; so, 20 was adopted as the operational span. Note that a larger span could have been adopted.

The F-scale is reversed to LO+20-F, as shown in figure 11, to associate the largest value of the sigmoidal span, 20, with the peak of the sigmoid at LO. Scaling control points from the left cross-section at S = 1240 (fig. 11) to 1.0 in X, (LO+20-F), and Y, (score, t), and making an overlay curve (see Matchacurve-1), an appropriate sigmoid was identified from the Standards.

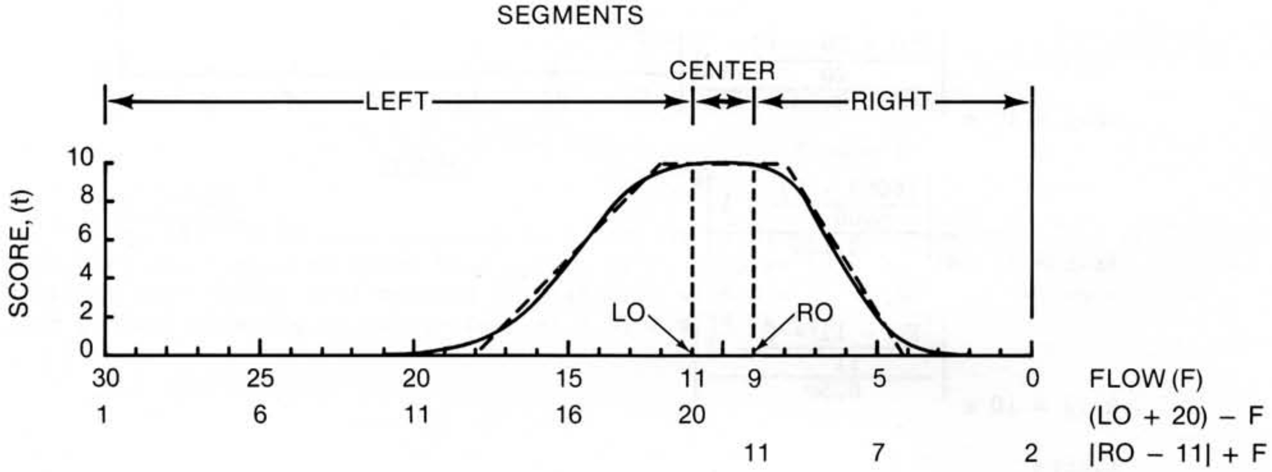

Figure 11.--t-cross-section over F at S = 5000. Note: The cross-sectional shape of the left segment is constant in the range 1240 ≤ S ≤ 5000; the right segment, although different in cross-sectional shape, is also constant in shape over the range 1240 ≤ S ≤ 5000.

Tsig

From figure 4, the planar ridge truncation ranges from t = 0 at S = 840, to t = 10 at S = 1240. This plane, scaled to 1.0, is matched and smoothed by Tsig, as shown in figure 12. Note from the Tsig formula specified (p. 13) that the inverted sigmoid to the left of the truncation plane (with base = 1.0 and peaking at 0, 0) was described on the reversed S-axis (6000-S) and subtracted from 1.0 to arrive at Tsig. This provided a more accurate duplication of the truncator plane than did other sigmoid alternatives, in this case.

Although a maximum of S = 6000 appears in Tsig, the applicable range is still limited to S = 5000 based on the original figure (fig. 4).

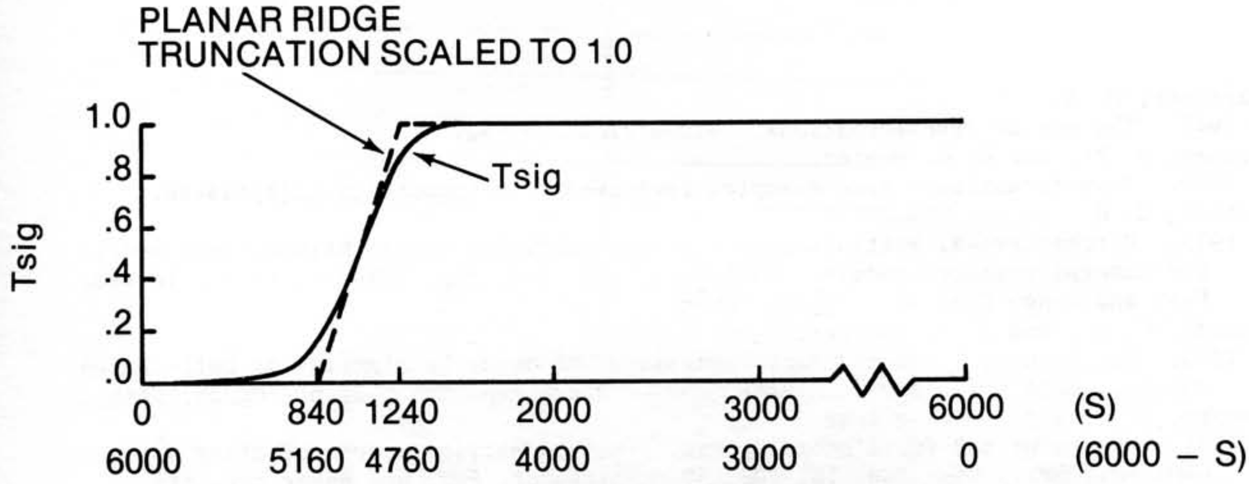

Figure 12.--Ridge truncation sigmoid, Tsig.

Rsig

The right segment of figure 4 has a constant cross-section over F for 1240 ≤ S ≤ 5000, approximated in the descriptor by a sigmoid oriented at RO and extending to the right for a distance of 11 units in F. Eleven is about the minimum operational span for the right-segment sigmoids. RO ranges from 8.404 at S = 0 to 10.808 at S = 5000. A range of 11 includes F = zero at both extremes, and so was adopted. The Rsigs were described as a function of the F-transform |RO-11|+F...as shown in figure 11. Thus, the maximum value of 11 always occurred at the sigmoidal peak, F = RO, and Rsig functioned over the range zero to 11 of the F-transform.

The final surface, then, is simply the independent sum of the three contiguous ridge segments...all truncated at appropriate points in S through multiplication by the proportional values of Tsig.

PUBLICATIONS CITED

Bartlett, M. S.
1947. The use of transformations. Biometrics 3:39-52.

Draper, N. R., and W. C. Hunter.
1969. Transformations: some examples revisited. Technometrics 11(1):23-40.

Jensen, C. E.
1973. Matchacurve-3, multiple-component and multidimensional mathematical models for natural resource models. USDA For. Serv. Res. Pap. INT-146, 42 p. Intermt. For. and Range Exp. Stn., Ogden, Utah.

Jensen, C. E., and J. W. Homeyer.
1970. Matchacurve-1 for algebraic transforms to describe sigmoid- or bell-shaped curves. USDA For. Serv., Intermt. For. & Range Exp. Stn., 22 p. Ogden, Utah

Jensen, C. E. and J. W. Homeyer.
1971. Matchacurve-2 for algebraic transforms to describe curves of class X^n. USDA For. Serv. Res. Pap. INT-106, 39 p. Intermt. For. and Range Exp. Stn., Ogden, Utah

Lee, D-Y, David, H. T., Mensing, R. W.
1973. Evaluation of gap-graded asphalt concrete mixtures. Part 1: Mechanical properties. Eng. Res. Inst., Iowa State Univ.

☆ U.S.GOVERNMENT PRINTING OFFICE:1976-O-677-328/51