1984

Development of Structured Regression Hypotheses Interactive/Descriptive Geometry Through Five Dimensions

United States Department of Agriculture, Forest Service

Follow this and additional works at: https://digitalcommons.usu.edu/govdocs_misc

Part of the Other Physical Sciences and Mathematics Commons

Recommended Citation

This Report is brought to you for free and open access by the U.S. Government Documents (Utah Regional Depository) at DigitalCommons@USU. It has been accepted for inclusion in Miscellaney by an authorized administrator of DigitalCommons@USU. For more information, please contact rebecca.nelson@usu.edu.
Development of Structured Regression Hypotheses / Interactive Descriptive Geometry Through Five Dimensions

Chester E. Jensen

THE AUTHOR
CHESTER E. JENSEN, retired, served as principal statistician for the Intermountain Forest and Range Experiment Station from 1967-80. He held the same position at the Northeastern and Central States Forest Experiment Stations prior to coming to the Intermountain Station. Mr. Jensen can be reached by mail through the Intermountain Station or by telephone at 801-621-0791.

ACKNOWLEDGMENTS
A large vote of thanks for the multitude of "free data" examples used in the text is due the author-scientists noted in each case. Their professional inputs were fundamental to compilation of the models generated.

My profound appreciation is extended also, to many scientists who graciously provided critical reviews of the past methods manuscripts summarized here, Dr. F. Thomas Lloyd, statistician (U.S. Forest Service, Southeastern Station), was perhaps the most consistent contributor in this respect.

RESEARCH SUMMARY
Regression hypotheses admitted to statistical evaluation should, of course, be germane to relations between the variables considered. This constraint is easily violated with the screening of arbitrarily adopted sets of hypotheses. Sequential polynomials in additive models, for example, may include components that are not commensurate with expectation (prior knowledge) and may not include those that are (e.g., unique interaction terms) and, since the analytical results from the screening of meaningless hypotheses may themselves be meaningless, there is a substantial incentive for serious prescreening consideration of the hypotheses to be evaluated.

Where prior knowledge is reasonably strong, the relation under consideration is highly interactive and/or has unique curve form or intercept constraints, the development of a fairly specific (structured) hypothesis is probably advisable. This can be evaluated on pertinent data sets not involved in the hypothesis development.

Prior knowledge information sources can include theory resulting from past investigation as well as information derived from data sets or parts of such data sets devoted to hypothesis development. Methods for developing multidimensional graphic models from two-dimensional effects found in partitioned data and for incorporating even complex constraints are shown in both contrived-and-live application examples.

Mathematical descriptors for the resulting graphic models are assembled using two families of curves, \( x^n \), \( n \geq 1 \); and \( e^x \), a highly manipulable function for bell-shaped curves or portions thereof. In either case, the X-axis for each two-dimensional effect is altered as needed to orient the curve in the upper right quadrant. The analyst then simply compares a normalized-scale version of the resulting curve with graphed standard curves.

Standards, and locates the two adjacent ones most nearly like his own in shape. Interpolation between the curves and their associated parameters, results in identification of the function for the best matching curve offered by the system.

The alternative of using a completely automated \( X \)-transform selection system from the Standards was considered at length. But, inability to define criteria (including the fitting process) that always result in the best choice of descriptor, regardless of data distribution or analytical objectives, resulted in retention of the less sophisticated, but still completely flexible graphic comparison technique shown herein. Methods of use are fairly well documented for each of the graphed-relation examples, which vary from two to five dimensions.

The mathematical descriptors can be regarded as pertinent hypotheses to be evaluated on data sets not used in their development. The evaluation on a single data set can involve a simple scale adjustment in \( Y \) of the hypotheses, as shown herein, or a variety of more complex testing formats (not shown) depending on complexity of the hypothesis itself, capacity of the analyst to linearize it, and/or on the nature of the fitting system adopted (linear, nonlinear). Also evaluation can be based on performance statistics compiled from real-world applications of the hypothesized descriptors.

Experience gained from the presentation of a dozen short courses covering methodology through the fourth dimension, suggests that this material can be assimilated by researchers with only college algebra-level mathemat-ical background and brief training in the operation of programmable electronic calculators. Also note that the graphic presentation scheme for the four-and-five-dimensional relations used in this paper can, with nominal explanation, be readily interpreted by land managers having little or no interest in the complex mathematical descriptors that may be involved.

TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
</tr>
<tr>
<td>Hypothesis Development - General</td>
</tr>
<tr>
<td>Graphic Hypothesis Development - Two Dimensions</td>
</tr>
<tr>
<td>Graphic Hypothesis Development - Four or More Dimensions</td>
</tr>
<tr>
<td>Mathematical Description of Graphed Hypotheses - Two Dimensions</td>
</tr>
<tr>
<td>Mathematical Description of Graphed Hypotheses - Four Dimensions</td>
</tr>
<tr>
<td>Publications Cited</td>
</tr>
<tr>
<td>Additional References</td>
</tr>
<tr>
<td>Appendix - A: eY Derivation</td>
</tr>
<tr>
<td>Appendix - B: ( X^n ) Standards</td>
</tr>
<tr>
<td>Appendix - C: eY Standards</td>
</tr>
</tbody>
</table>
Development of Structured Regression Hypotheses/Interactive Descriptive Geometry Through Five Dimensions

Chester E. Jensen

INTRODUCTION

Applications of statistics reported in biological literature were found by Mead and Pike (1975) to contain some prevalent and potentially damaging oversights. Among other things, they noted that, when transformations were used, simple polynomials predominated and seemed to be used as the simplest readily available smoothing curve, without any appeal to their theoretical properties as approximations to the true response function. From this, it appears that analysts are inclined to adopt polynomials as a convenient curve-form panacea for all regression hypotheses rather than undertake purposeful preanalysis development of hypotheses specific to the relation under study.

In the presumed presence of incentive for analysts to follow technical direction, these findings suggest that this is the current (1975, at least) state of regression art imo as applied statistics is concerned. One need only consult related text material to verify the notion. Hypothesis development is simply not a topic covered. Rather, integral polynomials are usually presented as a package of ready-made hypotheses that can be explored sequentially and mechanically using least squares fitting procedure. Discussion covering the inclusion of known constraints in these hypotheses is generally limited to that of the simple intercept. It is seldom, if ever, mentioned that the polynomials, or log-transforms, etc., cannot be presumed to provide adequate curve-form representation for all hypotheses. And, it is pertinent to note that failure to discuss hypothesis development and its potentially complex outcomes in text material constitutes de facto disapproval of any but simpler hypotheses, such as those by polynomials.

All told, there is ample cause for the hypothesis-related problems identified by Mead and Pike (1975) and attributed to authors of biological research papers. It would seem, however, that the source of the problem might more nearly be associated with limitations in text material and statistical curricula than with oversights by users of currently peddled statistical methodology.

An effort is made in this paper to initiate solutions to the hypothesis development problem through discussion and by showing analysts how to develop preanalysis hypotheses graphically, regardless of complexity, from accumulated knowledge and existing data. For biological subject matter, these hypotheses will be curvilinear and highly interactive as often as not.

Mathematical formulation of graphed hypotheses is accomplished through use of two families of curves: $X^n$, $n \geq 1$ (reprinted herein from Jensen and Homeroy 1970); and $e^{kX}$ (Jensen 1979), a highly manipulable function for bell-shaped curves or parts thereof. Descriptor development can be accomplished by analysts with a knowledge of basic algebra and access to a programmable calculator.

HYPOTHESIS DEVELOPMENT, GENERAL

A condition for valid statistical evaluation of hypotheses is that they be developed independently of data used for evaluation. So, the analyst's first undertaking in a regression analysis should be to establish the strongest possible concept of the underlying form of the relation from accumulated knowledge on the relations being investigated. It can be anticipated for a broad array of problems that these concepts will vary from extremely weak, with little or no basis for form definition, to very firm expectations for even specific algebraic forms.

In the weak-concept case, there is little reason to allow for any but simple curve-form hypotheses. Low-order polynomials in a linear model will probably suffice. Where well-established algebraic forms already exist for the form of the relation, the exact hypothesis is already available and evaluation can be initiated. The more common and perhaps more challenging circumstance, however, involves some intermediate level of prior knowledge where the analyst is sure of some elements in the relation being considered and less sure of others. His or her deliberations on the nature of the relation may result directly in a mathematical hypothesis or, more commonly, they may result in a graphed form wherein the analyst imposes reasonable intercept-, form-, and scale-contraints. These elements of the resulting graph, then, must be represented with suitable mathematical accuracy to establish the final intermediate-strength hypothesis.

If the graphed hypothesis involves unique curve forms, substantial effort might be required to search through existing functions like those in the list provided by Mead and Pike (1975), for those with the capacity to mimic unique forms of the hypothesis. (Note that mathematical descriptors should meet the acceptance criteria of the analyst involved [Bartlett 1947; Draper and Hunter 1969].)

A generally more efficient alternative may be that of analyst-developed functions to describe graphed hypotheses. Relatively simple methods for describing even complex multidimensional curvilinear interactions are included in this paper. Curves of only two classes are utilized: $X^n$, $n \geq 1$; and the very versatile three-parameter bell-shaped function $e^{kX}$. Methods and related examples shown include a variety of research applications to lend the realism to the instruction. Additional research applications not cited in the text are given under Additional References to provide the reader with the opportunity for expanded self-study of descriptor development.

Given a mathematical hypothesis, it can be evaluated using data sets not used in its development. The evaluation can vary from a test of general performance in $y = f(b)$ (hypothesis), as shown herein, to more complex formats designated to test performance of individual components of the hypothesis. The testing format adopted is likely to depend on the complexity of the hypothesis itself, the capacity of the analysts to linearize it, and/or on the nature of the fitting system to be used (linear, non-linear). Also evaluation can be based on performance statistics compiled over repeated real-world applications of the hypothesized descriptor. See Mosteller and Tukey (1977) and Box, Hunter, and Hunter (1978) for discussion of empirical and structural (mechanical) models, and fitting systems, etc.
GRAPHIC HYPOTHESIS DEVELOPMENT - TWO DIMENSIONS

Accumulated knowledge on the general nature of the regression relation under consideration may be used to enhance the derivation of finite graphed hypotheses from data. A common analytical plot is to dedicate part (perhaps half) of a current data set to the development of a hypothesis, reserving the rest for hypothesis evaluation.

To illustrate, accumulated knowledge regarding the relation between \( Y \) and \( X \) is such that the curve is expected to pass through the origin, to have a positive slope over \( X \), and to be linear-to-concave-upward in form. Within these constraints, a specific estimate of the true shape of the relation is developed from half of a current data set (fig. 1).

Here, the analyst visually estimates the curve form expressed in the data and simultaneously draws and hand-fits the form to the data by approximate least deviations (Karst 1958). All of this is accomplished within the constraints of expectation.

In another case, prior knowledge leads to a strong expectation for a positive intercept and, above this, a sigmoid upward to the right with \( Y \) asymptoting at a value of 100 percent over large values of \( X \). Within these constraints, a specific estimate is again developed from half of a current data set (fig. 2).

The foregoing procedures apply to any curve shape in any quadrant or combinations of quadrants. They also apply to two-dimensional segments within three-dimensional relations.

GRAPHIC HYPOTHESIS DEVELOPMENT - THREE DIMENSIONS

Extension of two-dimensional procedures to a three-dimensional case is shown in figure 3. Note that \( X_1 \) and \( X_2 \) are invariably correlated to a greater or lesser extent so that estimates of the \( X_1 \) effect will generally improve with a reduction in breadth of the \( X_2 \) segments. While only three \( X_2 \) segments are shown in figure 3 for the sake of simplicity, five or more segments would have been desirable if more data points had been available. The same procedure applies to additional independent variables, such as \( X_3 \) in figure 4 (following).

GRAPHIC HYPOTHESIS DEVELOPMENT - FOUR OR MORE DIMENSIONS

Extension of three-dimensional procedures to a four-dimensional case is shown in figure 4. The same procedures can be extended to five or more dimensions.
MATHEMATICAL DESCRIPTIONS OF GRAPHE D HYPOTHESES - TWO DIMENSIONS

Regardless of the number of dimensions and complexity of the relation to be described, it is ultimately sorted into two-dimensional effects that are described individually. These are finally reassembled, additively or multiplicatively as appropriate, to arrive at the whole descriptor.

Each two-dimensional element is described with either one or both of two families of curves, $A^n$, $n \geq 1$, or $e^{-k}$, a three-parameter sigmoidal function designed for easy use by the modeler.

Outline for "TWO DIMENSIONS"

**Class $X^n$, $n \geq 1$ (Exponential Curves).................................................. 10**
Curves flat-to concave-upward, positive slope:
In the right half of two-space (example 1) ............................................. 10
without constraint in two-space (example 2) ....................................... 14

Curves flat-to concave-upward, negative slope:
In the right half of two-space (example 3) ............................................. 15
without constraint in two-space (example 4) ....................................... 16

**Class $e^{-k}$ (Bell-Shaped or Sigmoidal Curves) ..................................... 21**
The implied curve fully expressible in the upper-right quadrant of two-space 21

Curves flat-to convex-upward, positive slope:
In the right half of two-space (example 5) ............................................. 17
without constraint in two-space (example 6) ....................................... 18

Curves flat-to convex-upward, negative slope:
In the right half of two-space (example 7) ............................................. 19
without constraint in two-space (example 8) ....................................... 20
MATHEMATICAL DESCRIPTIONS OF GRAPHTED HYPOTHESES – TWO DIMENSIONS

Alternative Arrangements in Space ........................................... 25

Using Applicable Parts of $e^x$ Curves ........................................ 28

Notes on $e^{-x}$ ................................................................. 29

Multiple-Component Descriptors ................................................ 29

Two-Component $X^n$, no Inflection ......................................... 33

Two-Component $X^n$, with Inflection ....................................... 33

Three-Component, $X^n$, with Flat Central Segment ..................... 34

Segmentation with $X^n$ .......................................................... 34

Straight-Line Slopes ($X^n$, $n = 1$) Plus or Minus Sigmoids .......... 35
MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES - TWO DIMENSIONS

Class- $X^n$, n $\geq$ 1, (Exponential Curves)

Reference is made to Matchacurve-2 (Jensen and Homeyer 1971) in which use of $X^n$ covers not only $n \geq 1$, but $0 < n \leq 1$ and $n < 0$ as well. Our experience since that work was published has indicated that almost all descriptor needs involving $X^n$ are reasonably well met with the ($n \geq 1$) option; so methods and Standards in this summary publication include the ($n \geq 1$) option only, a great simplification over the Matchacurve-2 presentation (Jensen and Homeyer, 1971).

Descriptor development for graphed hypotheses accommodates all possible orientations in the upper-right quadrant of two-dimensional space (Fig. 5), and those involving any combination of those up to four quadrants.

Regardless of the form or quadrant(s) involved, the X-axis can be altered as needed to orient the curve in the upper-right quadrant and through the origin. The analyst simply compares a normalized-scale version of this reoriented hypothesis with graphed standard curves, or Standards (appendix, marked for cut-out), and selects the two adjacent ones most nearly like his own in shape. Interpolation between the curves and their associated powers ($n$) results in identification of the best alternative offered by the system.

Note that if duplicates are made of the standards, the copies should be exactly the same scale as the Standards.

Example 1
Here (Fig. 6), the hypothesis is a flat-to-concave-upward, positive slope, and occupying any vertical position in the right half of two-dimensional space. Mathematical formulation for the middle alternative (zero Y-intercept) is detailed in figure 7 and in table 1, but the methods will accommodate any Y-intercept.

Figure 5.—Alternative orientations of $X^n$ curves in the upper-right quadrant.

Figure 6.—$Y = a + bX^n$, $0 < X < XP$, n $\geq 1$. 
In figure 7, the analyst visually estimates the curve form expressed in the data and simultaneously draws and hand-fits the form to the data by approximate least deviations. All of this is accomplished within the constraints of expectation. Mathematical formulation procedures follow in table 1.

Table 1.—Mathematical hypothesis development

<table>
<thead>
<tr>
<th>Control</th>
<th>X</th>
<th>Y</th>
<th>X1</th>
<th>Y1</th>
<th>XP1</th>
<th>YP1</th>
<th>Y1 - YP1</th>
<th>Y2</th>
<th>Y2 - YP2</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>points</td>
<td>X(1)</td>
<td>Y(2)</td>
<td>X1(3)</td>
<td>Y1(4)</td>
<td>XP1(5)</td>
<td>YP1(6)</td>
<td>Y1 - YP1(7)</td>
<td>Y2(8)</td>
<td>Y2 - YP2(9)</td>
<td>Y(11)</td>
</tr>
<tr>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>200</td>
<td>1.3</td>
<td>200</td>
<td>1.3</td>
<td>25</td>
<td>52</td>
<td>1.6</td>
<td>3</td>
<td>1.7</td>
<td>4</td>
<td>1.7</td>
</tr>
<tr>
<td>400</td>
<td>10.6</td>
<td>400</td>
<td>10.6</td>
<td>.50</td>
<td>.16</td>
<td>10.2</td>
<td>-4</td>
<td>10.8</td>
<td>0</td>
<td>10.6</td>
</tr>
<tr>
<td>600</td>
<td>31.4</td>
<td>600</td>
<td>31.4</td>
<td>75</td>
<td>48</td>
<td>30.4</td>
<td>-10</td>
<td>30.9</td>
<td>-5</td>
<td>30.9</td>
</tr>
<tr>
<td>800</td>
<td>66.0</td>
<td>800</td>
<td>66.0</td>
<td>1.00</td>
<td>1.00</td>
<td>66.0</td>
<td>0</td>
<td>66.0</td>
<td>0</td>
<td>66.0</td>
</tr>
</tbody>
</table>

- Columns 1 and 2: Select representative control points (point coordinates) from the graphed hypothesis, including XP, the largest value of X to be represented by the hypothesis (800, in this case), and YP, the corresponding smoothed value of Y (66 here). Five control points have been selected and are listed in these columns, along with computations necessary to the search for a suitably similar mathematical curve-form. Use fewer control points where the desired sensitivity is less; use more for greater sensitivity.
- Column 3: The X-scale remains unchanged, but is relabeled as X1 for consistency with procedures for subsequent, more complex examples. XP is called XP1 in the X1-scale.
- Column 4: Subtract the intercept, a, from all control point Y-values to produce Y1. Whether a is zero (as here) or any positive or negative value, Y1-values will have zero intercept and will be in the upper-right quadrant of two-dimensional space. Such orientation of the curve is necessary to use of the Standards. YP is called YP1 in the Y1-scale.
- Columns 5 and 6: Compute ratios of X1 and Y1 to XP1 and YP1, respectively. These arrays of ratios are the normalized (proportional) scales for the X1 and Y1 control points. The resulting overlay can now be plotted on a sheet of conventional graph paper (10 × 10 to 1/2 inch is best), at the exact scale specified in figure 8. Although not necessary, it may help to draw a smooth curve through the points. Call this the overlay curve. It is the normalized version of the original hypothesis plotted at the same scale used for the Standards (fig. 8).

Figure 7.—Y = 0 + bX2
0 ≤ X ≤ XP, n ≥ 1.

Figure 8.—Here the graphed curve has been normalized and plotted at the same scale as the Standards.

- Place the overlay graph over the Standards, being certain that the X- and Y-axes are matched exactly. Find the Standards that most nearly bracket the overlay curve (fig. 9). In this case, n = 2.5 and 3.0 bracket the overlay curve nicely. Note that a light source behind the graphs is being compared who assist in the matching process. Where the hypoth-
ized curve is not suitably matched by the Standards, a segmented or multicomponent Xn-function, or some portion of a sigmoid may suffice (see Contents).

Use proportional departure of the overlay curve from the left bracketing Standard to approximate an interpolated value between n = 2.5 and 3.0. This is best done at a point where curve-form accuracy is most crucial—the point of sharpest bend was adopted in this case. From figure 9, the interpolated n is 2.7.
- Column 7: After the analyst has made the first approximation of the mathematical form of the hypothesis, Y1 = hXYP1, where h = YP1/XP1 = 66/(900) = 7.36 x 10⁻², he can make associated estimates, Y1, at each control point XP1 value as shown.
- Column 8: Differences between estimated and original Y1 values indicate that hXYP1 has a preponderance of negative departures (that is, too much curvature). Try flatter curves (less power) iteratively.
MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES - TWO DIMENSIONS

**Figure 9.** The Standards best matched to the original curve bracket the overlay curve.

- Column 9: A second estimator, $Y_2 = b(X)^{2.54}$, where $b = YP(XP)^{2.54} = 1.430 \times 10^{-6}$, produces $Y_2$ values as shown.
- Column 10: Differences between estimated $Y_2$ and original $Y_1$ values are small and well balanced (+ and -); so $b(X)^{2.54}$ is adopted.
- Column 11: The final hypothesis is then, $Y = a + bX = 0 + 1.430 \times 10^{-6}(X)^{2.54}$, or, since $X=X_1$, $Y = 1.430 \times 10^{-6}(X)^{2.54}$.

![Graph](image)

**Figure 10.** The fitted model has been plotted over the original graphed form.

Note that at the scale shown in figure 10, differences between the graphed hypothesis and $Y = 1.430 \times 10^{-6}(X)^{2.54}$ are barely discernible. This example produced a closer mathematical match for the objective curve than one might expect on the average, but serves to demonstrate the procedures very well.

Table 2. Mathematical hypothesis development

<table>
<thead>
<tr>
<th>Control points</th>
<th>$(X - k)$</th>
<th>$(Y - a)$</th>
<th>Overlay control points</th>
<th>$b(X)^{2.54}$</th>
<th>Performance of $Y_1$</th>
<th>Performance of $Y_2$</th>
<th>$a + Y_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>$Y$</td>
<td>$X_1$</td>
<td>$Y_1$</td>
<td>$X_1P(X)$</td>
<td>$Y_1P(Y)$</td>
<td>$(Y_1 - Y)$</td>
<td>$Y_2$</td>
</tr>
<tr>
<td>..................</td>
<td>...........</td>
<td>.........</td>
<td>.........</td>
<td>...........</td>
<td>...........</td>
<td>...............</td>
<td>.......</td>
</tr>
<tr>
<td>400</td>
<td>-10.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>200</td>
<td>-8.7</td>
<td>200</td>
<td>1.3</td>
<td>25</td>
<td>0.02</td>
<td>1.6</td>
<td>+3</td>
</tr>
<tr>
<td>0</td>
<td>400</td>
<td>10.6</td>
<td>-40</td>
<td>10.2</td>
<td>-4</td>
<td>10.6</td>
<td>0</td>
</tr>
<tr>
<td>200</td>
<td>21.4</td>
<td>600</td>
<td>31.4</td>
<td>75</td>
<td>46</td>
<td>30.4</td>
<td>-1.0</td>
</tr>
<tr>
<td>400</td>
<td>56.0</td>
<td>800</td>
<td>96.0</td>
<td>1.00</td>
<td>1.00</td>
<td>1.00</td>
<td>66.0</td>
</tr>
</tbody>
</table>

**Figure 11.** $Y = a + b(X - k)^{2}$, $k = k; X; X, n = 1$.
MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES - TWO DIMENSIONS

• Columns 1 and 2: List control points representative of the curve.
• Shift the curve to the upper-right quadrant to match the position of the Standards:
  Column 3: Let \( X_1 = (X - k) \), where \( k \) is the smallest value of \( X \) to be represented by the equation (-400 here). Then, \( X_1 = (X - [-400]) = (X - 400) \), and \( XPI = 800 \).
• Column 4: \( Y_1 = (Y - a) - (Y - [10]) = (Y - 10) \), and \( YPI = 66.0 \).
• Columns 5-10: Identical to example 1, table 2, columns 5-10.
• Column 11: The final hypothesis is then, \( Y = a + bX^n = -10 + 1.430X/10 + (X - 400)^2/4 \), 
  \(-400 \leq X \leq 400.\)
• Least squares fit.
  With \( y \) fixed; \( Y = \exp(XT) \), where \( XT = b(X1) \), \( c = \exp(XT + Y1) \), \( XT^2 \), and \( Y - a = \exp(X1) \).
  With \( y \) variable, \( Y = a + b(X1) \). Letting \( X = (X1) \) and, under conventional notation, \( b = \exp(X1) \) and \( a = bX \).

Example 3
Assume a hypothesis like that in figure 12 (flat-to-concave-upward, negative slope, and occupying any vertical position in the right half of two-dimensional space). Mathematical formulation for the middle alternative (\( a = 0 \)) is detailed in table 3, but the methods will accommodate any \( a \) value. Procedures follow closely those of example 1.

Table 3—Mathematical hypothesis development

<table>
<thead>
<tr>
<th>Control points</th>
<th>( XP - X )</th>
<th>( Y - a )</th>
<th>Overlay control points</th>
<th>( b(X1)^2 )</th>
<th>Performance of ( Y1 )</th>
<th>Performance of ( Y2 )</th>
<th>Performance of ( Y )</th>
<th>( a + Y2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X )</td>
<td>( Y )</td>
<td>( X1 )</td>
<td>( Y1 )</td>
<td>( X1P1 )</td>
<td>( Y1P1 )</td>
<td>( Y1 )</td>
<td>( Y1 - Y )</td>
<td>( Y2 )</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td>0</td>
<td>66.0</td>
<td>500</td>
<td>66.0</td>
<td>1.00</td>
<td>1.00</td>
<td>66.0</td>
<td>0.0</td>
<td>66.0</td>
</tr>
<tr>
<td>200</td>
<td>31.4</td>
<td>600</td>
<td>31.4</td>
<td>.75</td>
<td>.48</td>
<td>30.4</td>
<td>-1.0</td>
<td>30.9</td>
</tr>
<tr>
<td>400</td>
<td>10.6</td>
<td>400</td>
<td>10.6</td>
<td>.50</td>
<td>.16</td>
<td>10.2</td>
<td>-4</td>
<td>10.6</td>
</tr>
<tr>
<td>600</td>
<td>1.3</td>
<td>200</td>
<td>1.3</td>
<td>.35</td>
<td>.02</td>
<td>1.6</td>
<td>+3</td>
<td>1.7</td>
</tr>
<tr>
<td>800</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0.0</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES - TWO DIMENSIONS

• Columns 1 and 2: List control points representative of the curve.
• Shift the curve to the upper-right quadrant to match the position of the Standards:
  Column 3: Let \( X_1 = (XP - X) = (800 - X) \), and \( XPI = 800 \).
• Column 4: Since \( a = 0 \) zero here, \( Y_1 = (Y - a) = (Y - 0) = Y \), and \( YPI = 66.0 \).
• Columns 5-10: Identical to example 1, table 2, columns 5-10.
• Column 11: The final hypothesis is then, \( Y = a + b(XP - X)^r = 0 + 0.1430 \times (800 - X)^2/4, \) \( 0 \leq X \leq 800.\)
• Least squares fit.
  With \( y \) fixed at any value; \( Y = \exp(XT) \), where \( XT = b(X1) \), \( c = \exp(XT + Y1) \), \( XT^2 \), and \( Y - a = \exp(X1) \).
  With \( y \) variable; \( Y = a + b(X1) \). Letting \( X = (X1) \) and, under conventional notation, \( b = \exp(X1) \) and \( a = bX \).

Example 4
Assume a hypothesis like that in figure 13 (flat-to-concave-upward, negative slope, and crossing into the left half of (two-dimensional space)). Mathematical formulation procedures follow closely, those of example 1 and are summarized in table 4.

Table 4—Mathematical hypothesis development

<table>
<thead>
<tr>
<th>Control points</th>
<th>( XP - X )</th>
<th>( Y - a )</th>
<th>Overlay control points</th>
<th>( b(X1)^2 )</th>
<th>Performance of ( Y1 )</th>
<th>Performance of ( Y2 )</th>
<th>Performance of ( Y )</th>
<th>( a + Y2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X )</td>
<td>( Y )</td>
<td>( X1 )</td>
<td>( Y1 )</td>
<td>( X1P1 )</td>
<td>( Y1P1 )</td>
<td>( Y1 )</td>
<td>( Y1 - Y )</td>
<td>( Y2 )</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td>-400</td>
<td>56.0</td>
<td>800</td>
<td>66.0</td>
<td>1.00</td>
<td>1.00</td>
<td>66.0</td>
<td>0.0</td>
<td>66.0</td>
</tr>
<tr>
<td>-200</td>
<td>21.4</td>
<td>600</td>
<td>31.4</td>
<td>.75</td>
<td>.48</td>
<td>30.4</td>
<td>-1.0</td>
<td>30.9</td>
</tr>
<tr>
<td>0</td>
<td>6</td>
<td>400</td>
<td>10.6</td>
<td>.50</td>
<td>.16</td>
<td>10.2</td>
<td>-4</td>
<td>10.6</td>
</tr>
<tr>
<td>800</td>
<td>1.3</td>
<td>200</td>
<td>1.3</td>
<td>.35</td>
<td>.02</td>
<td>1.6</td>
<td>+3</td>
<td>1.7</td>
</tr>
<tr>
<td>800</td>
<td>-10.0</td>
<td>0</td>
<td>0</td>
<td>0.00</td>
<td>0.0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

BEST COPY AVAILABLE

BEST COPY AVAILABLE
MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES – TWO DIMENSIONS

- Columns 1 and 2: List control points representative of the curve.
- Shift the curve to the upper-right quadrant to match the position of the Standards:
  Column 3: Let \( X_1 = (X_P - X) = (400 - X) \), and \( X PI = 800 \).
  Column 4: Let \( Y_1 = (Y - a) + (Y' - (-10)) = (Y + 10) \), and \( Y PI = 66.0 \).
- Column 5-10: Identical to example 1, table 2, columns 5-10.
- Column 11: The final hypothesis is then,
  \[ Y = a + b(X_P - X')^p = -10 + 1.430 \times 10^{-5}(400 - X)^{1.34}, \]
  \( k \leq X \leq 400 \), and \( k \) is the lower applicable limit (400 here) of the \( X \)-range.
- Least squares fit.

With a fixed at any value: \( Y = c(XT) \), where \( XT = b(X1)^p \), \( c = \Sigma(XT)^2 / \Sigma(XT)^2 \), and \( Y = a + c(b(X1))^p \).

With a variable, \( Y = a + b(X1)^p \). Letting \( X = (X1)^p \) and, under conventional notation, \( b = \Sigma(XT)^2 \) and \( c = \Sigma(XT) \).

Example 5

Assume a hypothesis like that in figure 14 (flat-to convex-upward, positive slope, and occupying any vertical position in the right half of two-dimensional space). Mathematical formulation for the middle alternative is detailed in table 5. Procedures follow closely those of example 1.

![Figure 14](image)

**Table 5.** Mathematical hypothesis development

<table>
<thead>
<tr>
<th>Control points</th>
<th>( (XP - X) )</th>
<th>( (YP - Y) )</th>
<th>Overlay control points</th>
<th>( b(XT)^{1.34} )</th>
<th>Performance of ( b(XT)^{1.34} )</th>
<th>Performance of ( b(XT)^{1.34} )</th>
<th>( p + Y2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X )</td>
<td>( Y )</td>
<td>( X1 )</td>
<td>( Y1 )</td>
<td>( X1/XP )</td>
<td>( Y1/YP )</td>
<td>( Y1 (Y1 - Y) )</td>
<td>( Y2 (Y2 - Y) )</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
</tr>
<tr>
<td>0</td>
<td>0.0</td>
<td>800</td>
<td>66.0</td>
<td>1.00</td>
<td>1.00</td>
<td>66.0</td>
<td>66.0</td>
</tr>
<tr>
<td>200</td>
<td>34.6</td>
<td>600</td>
<td>31.4</td>
<td>.75</td>
<td>.75</td>
<td>30.4</td>
<td>-10</td>
</tr>
<tr>
<td>400</td>
<td>55.4</td>
<td>400</td>
<td>10.2</td>
<td>.50</td>
<td>.50</td>
<td>10.2</td>
<td>-4</td>
</tr>
<tr>
<td>600</td>
<td>64.7</td>
<td>200</td>
<td>1.3</td>
<td>.25</td>
<td>.25</td>
<td>1.6</td>
<td>+3</td>
</tr>
<tr>
<td>800</td>
<td>66.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

![Figure 15](image)

**Example 6**

Assume a hypothesis like that in figure 15 (flat-to convex-upward, positive slope with negative \( Y \)-intercept, and crossing into the left half of two-dimensional space). Mathematical formulation for this curve is detailed in table 6. Procedures closely follow those of example 1.

**Table 6.** Mathematical hypothesis development

<table>
<thead>
<tr>
<th>Control points</th>
<th>( (XP - X) )</th>
<th>( (YP - Y) )</th>
<th>Overlay control points</th>
<th>( b(XT)^{1.34} )</th>
<th>Performance of ( Y1 )</th>
<th>( b(XT)^{1.34} )</th>
<th>Performance of ( Y2 )</th>
<th>( a + Y2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X )</td>
<td>( Y )</td>
<td>( X1 )</td>
<td>( Y1 )</td>
<td>( X1/XP )</td>
<td>( Y1/YP )</td>
<td>( Y1 (Y1 - Y) )</td>
<td>( Y2 (Y2 - Y) )</td>
<td>( (Y2 - Y1) )</td>
</tr>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
<td>(8)</td>
<td>(9)</td>
</tr>
<tr>
<td>( -400 )</td>
<td>-40.0</td>
<td>800</td>
<td>66.0</td>
<td>1.00</td>
<td>1.00</td>
<td>66.0</td>
<td>66.0</td>
<td>0.0</td>
</tr>
<tr>
<td>( -200 )</td>
<td>-5.4</td>
<td>600</td>
<td>31.4</td>
<td>.75</td>
<td>.75</td>
<td>30.4</td>
<td>-10</td>
<td>30.9</td>
</tr>
<tr>
<td>0</td>
<td>0.0</td>
<td>400</td>
<td>10.6</td>
<td>.50</td>
<td>.50</td>
<td>10.2</td>
<td>-4</td>
<td>10.6</td>
</tr>
<tr>
<td>( 200 )</td>
<td>24.7</td>
<td>200</td>
<td>1.3</td>
<td>.25</td>
<td>.25</td>
<td>1.6</td>
<td>+3</td>
<td>1.7</td>
</tr>
<tr>
<td>( 400 )</td>
<td>66.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

**Figure 15.** \( Y = (XP)^{1.34} \)
Mathematical Descriptions of Graphed Hypotheses – Two Dimensions

- Columns 1 and 2: List control points representative of the curve.
- Shift the curve to the upper-right quadrant to match the position of the Standards:
  - Column 3: Let $X_1 = ( XP - X) = 400 - X$ and $XP = 800$.
  - Column 4: Let $Y_1 = (YP - Y) = 66.0 - Y$ and $YP = 66.0$.
- Columns 5-10: Identical to example 1, table 2, columns 5-10.
- Column 11: The final hypothesis is then,
  \[ Y = YP - b(XP - X)^p = 66.0 - 1.430 \times 10^{-0.3333} \times (-400 - X)^{1.25}, \]
  for $-400 \leq X \leq 400$.
- Least squares fit.
  With YP fixed at any value; $\hat{Y} = c(XT)$, where $XT = b(XY^*)$, $c = \sum(XX^*)/\sum(XX^T)$, and $\hat{Y} = YP - c(b(X)^*)$.
  With YP variable, $\tilde{Y} = a + b(X)^*$. Letting $X = (X)^*$ and, under conventional notation, $b = \Sigma xy/\Sigma x^2$, and $a = \tilde{Y} - b\tilde{x}$.

Example 7
Assume a hypothesis like that in figure 16 (flat to convex-upward, negative slope and occupying any vertical position in the right half of two-dimensional space). Mathematical formulation for the middle alternative is detailed in table 7. Procedures closely follow those of example 1.

![Figure 16](image)

Table 7—Mathematical hypothesis development

<table>
<thead>
<tr>
<th>Control points</th>
<th>X</th>
<th>(YP - Y)</th>
<th>Overlay control points</th>
<th>Overlay performance of Y</th>
<th>Performance of Y</th>
<th>$a + Y_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>Y</td>
<td>$X_1$</td>
<td>$Y_1$</td>
<td>$X_1(UP)$</td>
<td>$Y_1(UP)$</td>
<td>$\hat{Y}_1$</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>------</td>
<td>------</td>
<td>---------</td>
<td>---------</td>
<td>------</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>200</td>
<td>64.7</td>
<td>200</td>
<td>1.3</td>
<td>25</td>
<td>0.02</td>
<td>1.6</td>
</tr>
<tr>
<td>400</td>
<td>55.4</td>
<td>400</td>
<td>10.6</td>
<td>50</td>
<td>16</td>
<td>10.2</td>
</tr>
<tr>
<td>600</td>
<td>34.6</td>
<td>600</td>
<td>31.4</td>
<td>75</td>
<td>48</td>
<td>30.4</td>
</tr>
<tr>
<td>800</td>
<td>0</td>
<td>800</td>
<td>66.0</td>
<td>1.00</td>
<td>100</td>
<td>66.0</td>
</tr>
</tbody>
</table>

Example 8
Assume a hypothesis like that in figure 17 (flat to convex-upward, negative slope and crossing into the left half of two-dimensional space). Mathematical formulation for this curve is detailed in table 8. Procedures closely follow, those of example 1.

![Figure 17](image)

Table 8—Mathematical hypothesis development

| Control points | X | Y | X1 | Y1 | X1(UP) | Y1(UP) | $\hat{Y}_1$ | (Y1 - Y1) | Y2 | (Y2 - Y1) | $\hat{Y}$ |
|----------------|---|---|----|----|--------|--------|--------|--------|----|--------|------|--------|
| X | Y | X1 | Y1 | X1(UP) | Y1(UP) | $\hat{Y}_1$ | (Y1 - Y1) | Y2 | (Y2 - Y1) | $\hat{Y}$ |
|---|---|----|----|--------|--------|--------|--------|----|--------|------|--------|
| 0 | 0.0 | 0.0 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 0 | 15.4 | 400 | 10.6 | 50 | 16 | 10.2 | -4 | 10.6 | 0 | 55.4 |
| 200 | 14.7 | 200 | 1.3 | 25 | 0.02 | 1.6 | +3 | 1.7 | +4 | 64.3 |
| 400 | 10.6 | 400 | 10.6 | 50 | 16 | 10.2 | -4 | 10.6 | 0 | 55.4 |
| 600 | 7.5 | 600 | 31.4 | 75 | 48 | 30.4 | -10 | 30.9 | -5 | 35.1 |
| 800 | 0 | 800 | 66.0 | 1.00 | 100 | 66.0 | 0 | 66.0 | 0 | 0 |

- Columns 1 and 2: List control points representative of the curve.
- Shift the curve to the upper-right quadrant to match the position of the Standards:
MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES – TWO DIMENSIONS

Column 3: Let \( X = (X - k) \), where \( k \) is the smallest value of \( X \) to be represented by the equation, \(-400 \) here. Then, \( X_1 = (X + 400) \), and \( X_{P1} = 800 \).

Column 4: Let \( X = (Y_P - Y) \), where \( Y_P = 60 \).

* Column 5-10: Identical to example 1, table 2, columns 5-10.

* Column 11: The final hypothesis is then,
\[
\hat{Y} = Y - b(X)^p = 26.0 - 1.430 \times 10^{-4} (X)^4, \quad -400 \leq X \leq 400.
\]

* Least squares fit.
With \( Y_P \) fixed, at any value; \( \hat{Y} = c(X) \), where \( XT = b(X) \), \( c = \Sigma(XT)^n / \Sigma(XT)^n \), and \( \hat{Y} = Y - c(X) \).

With \( Y \) variable, \( \hat{Y} = a + b(X)^P \). Letting \( X = (X)^P \) and, under conventional notation, \( b = \Sigma(X) / \Sigma(X)^2 \), and \( a = \hat{Y} - bX \).

Class \( e^X \) (Bell-shaped or Sigmoidal Curves)
By reason of \( e^X \) symmetry, shape parameters for this bell-shaped function are identical for either sigmoidal half (see \( e^X \) ITS DERIVATION, CAPABILITIES, AND LIMITATIONS, appendix A). So, the \( e^X \) Standards are limited to sigmoids and these have normalized (proportional) \( X \) and \( Y \)-scales, in the upper right quadrant of two-dimensional space, and have zero intercept and slope upward to the right.

Representative control points (point-coordinates) are selected from a graphed sigmoidal hypothesis and these points are shifted to the position of the Standards through manipulation of the \( X \)-and \( Y \)-scales as needed. These revised scales are normalized (converted to proportions) and the normalized control points are plotted on graph paper at the exact scale of the Standards. This point-representation of the hypothesis (with or without the corresponding curve sketched in) can be laid over and compared with alternative sigmoids on any page of the Standards. When Standards are found that bracket the overlay curve and are suitably similar to it, the analyst can interpolate between parameters of these Standards to arrive at the best sigmoid provided by the system.

The result is a mathematical hypothesis that can be fitted to new data and evaluated statistically. For example, assume the graphed hypothesis shown in figure 18.

![Figure 18](image-url)

MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES – TWO DIMENSIONS

Recognizing that the shape of the curve on each side of the peak is about the same, it seems likely that a symmetrical curve fitting the left, most completely specified side of the relation, will also fit the top segment of the implied sigmoid to the right of the peak. So, we only need identify the mathematical form of the sigmoid to the left of the peak (0 ≤ X ≤ 150) to describe the entire graphed hypothesis. Mathematical formulation of this curve is detailed in table 9.

Table 9. Mathematical hypothesis development

<table>
<thead>
<tr>
<th>Control points</th>
<th>( X )</th>
<th>( Y )</th>
<th>Overlay control points</th>
<th>( n = 2 ) \times \text{ST} )</th>
<th>Performance of ( Y )</th>
<th>Performance of ( Y' )</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X )</td>
<td>( Y )</td>
<td>( X )</td>
<td>( Y )</td>
<td>( Y_{P1} = 100 )</td>
<td>( Y_{P1} = 100 )</td>
<td>( Y_{P1} = 100 )</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>75</td>
<td>100</td>
<td>75</td>
<td>100</td>
<td>50</td>
<td>10</td>
<td>100 ( \times 10 )</td>
</tr>
<tr>
<td>105</td>
<td>430</td>
<td>105</td>
<td>430</td>
<td>105</td>
<td>430</td>
<td>105</td>
</tr>
<tr>
<td>150</td>
<td>100</td>
<td>150</td>
<td>100</td>
<td>100</td>
<td>100</td>
<td>100</td>
</tr>
</tbody>
</table>

Procedures generally follow those of example 1 for curves of the \( e^X \)-class.

* Columns 1 and 2: Five control points representative of the curve (including \( XP \), the point in \( X \) at which \( Y \) peaks, and \( Y_{P} \), the peak value of \( Y \)) have been selected from the sigmoid and are listed here. Use fewer control points where the desired sensitivity is less, more for greater sensitivity.

* Columns 3 and 4: The Standards for \( e^X \) are oriented at zero in the upper-right quadrant (as for the \( e^X \) Standards) and, since this curve already conforms, there are no transformations of the \( X \)- and \( Y \)-axes. So, \( X \times 1 \) and \( Y \times 1 \), respectively. The \( X \)- and \( Y \)-variables will be shown and used here to be consistent with procedures for more complex examples that follow. Note that \( XP \) is the maximum value of \( X \) and that \( Y_{P1} \) is the maximum value of \( Y \).

* Columns 5 and 6: Compute ratios of \( X_{P} \) and \( Y_{P1} \) to \( X \) and \( Y \), respectively.

These arrays of ratios are the normalized (proportional) scales for the \( X \)- and \( Y \)-control points. The resulous overlay curves can now be plotted on a sheet of conventional graph paper (10 × 10 or ½ inch is best), at the exact scale specified in figure 19 (5 inches for the \( Y_P \)-range and ½ inch per 1/10 unit of \( Y \), 1½ inches for the \( XP \)-range and 1/4 inch per 1/10 unit of \( X \)). Although not necessary, it may help to sketch a smooth curve through these plotted points as shown in figure 19.

Use this curve as an overlay for any page of the Standards, being careful to match exactly, the \( X \)- and \( Y \)-axes of the overlay with those of each page examined. Find adjoining sigmoids that are most nearly shaped like the overlay curve and bracket it (fig. 20).

Note in the Standards that \( Is \) is the proportional distance of the inflection point from zero to 1.0 and \( n \) is the power of the negative exponents for \( e^X \) (appendix). Curvature of the sigmoids becomes more pronounced with increased \( n \) and changes between

![Figure 18](image-url)
Having found an acceptable pair of adjoining sigmoids that bracket the overlay curve, the analyst can interpolate between the proportional inflection points they represent and thus select a sigmoid more nearly matching the hypothesis. Proportional horizontal departure of the overlay curve from the left bracketing sigmoid at \( Y / Y_P = 0.5 \) can be used as the basis for interpolation. In figure 20, the overlay curve lies at about the 70 percent point between inflection points 0.6 and 0.7; so the interpolated \( f = 0.67 \). (This same interpolating principle can also be used between Standards of different sets for refined "n" estimates as the need arises.) Then substituting in the parent \( e^A \) equation, we have:

\[
\begin{align*}
Y_1 &= \frac{X - 1}{1 - e^{-X}} = \frac{X - 1}{1 - e^{-X}} \\
Y_2 &= \frac{X - 1}{1 - e^{-X}}
\end{align*}
\]

This is easily programmed on a small desk-top computer (with the number of steps required generally being less than 100) and solved for \( Y_1 \)-values at control-point \( X_1 \)-values.

- Column 7: \( e^A \) (ranging only from 0 to 1) is scaled to the graphed hypothesis by simply multiplying \( e^A \) by \( Y_P \) (\( = Y_P \) here), the height of the sigmoid at its peak. Adding the intercept \( a \), produces \( Y_1 \).

- Column 8: The selected sigmoid is reasonably close to that of the hypothesis with a slight surplus of positive values (+0.6), as shown. In the light of the \( Y \)-values involved (100, 430, 700), the departures +0.6, +7.5, and −7.4, respectively, are small and this sigmoid would most likely be adopted. But, as might be more appropriate in other cases, computer iteration can be used to select an even closer \( e^A \)-approximation of the hypothesized sigmoid.

- Columns 9 and 10: \( Y_2 \) and its performance are slightly better at \( f = 0.670 \). Assuming this to be the final form adopted, we have:

\[
Y_2 = 0 + 1000 \left( \frac{X - 1}{150} \right)^2 = 0.3299 \frac{X - 1}{1 - 0.001}
\]

And, since the right numerator is very small and the denominator approaches 1.0,

\[
Y_2 = 0 + 1000 e^{0.001} \left( \frac{X - 1}{150} \right)^2 = 0.3299 \frac{X - 1}{1 - 0.001}
\]

Note that at the scale shown in figure 21, differences between the graphed hypothesis and \( Y_2 \) are not discernible. This example produced a closer mathematical match for the objective curve than one might expect on the average, but serves to demonstrate the procedures very well.
matter of fitting methods interest, the hypothesis was then fitted by least squares (forced through the origin) back to the data from which it was partially derived.

And: $\hat{Y} = \hat{c}X_T$, $X_T = 1000*e^{(Y/X)\hat{c}/1000}$

almost no change.

which again, is not distinguishable from the hypothesis at the scale shown in figure 21.

While this result shows the development process has not deteriorated the fit of the hypothesis to the source data, statistical evaluation should be reserved for a new data set. The same fitting procedures are applicable.

Alternative Arrangements in Space

Unique spacial arrangement of the bell-shaped or sigmoidal form may necessitate adjustment of the $Y$-values for an intercept and/or manipulation of the $X$-scale to shift the objective figure to the upper right quadrant (like the Standards). Commonly encountered spacial alternatives and associated shifting techniques are shown in figures 22-26 which follow.
MATHEMATICAL DESCRIPTIONS OF GRAPHD HYPOTHESES - TWO DIMENSIONS

- Plot the mirror image of the right side over the range $0 \leq X \leq XP$ and identify an appropriate sigmoid using $X$ and $Y$-values from that range, as in table 9. The sigmoid adopted will apply equally on either side of $XP$.

![Figure 24](image)

Figure 24.—$X = 0$ included in the range $XP \times XP$.

- Substitute $X1 = (X - k)$ for $X$ in the range $XP \leq X \leq (XP + k)$. Then follow table 9 procedures to identify an $e^{-x}$-sigmoid for the left side of the bell-shaped curve ($=$ right side by virtue of symmetry here.)

![Figure 25](image)

Figure 25.—Incomplete sigmoid specified.

- Complete the implied sigmoid in accord with expectation and identify $e^{-x}$ as in table 9. The hypothesis would be presumed to apply in the range $0 \leq X \leq XP$.

![Figure 26](image)

Figure 26.—Inverted bell-shaped curve, or portion thereof.

Using Applicable parts of $e^{x}$-curves

On occasion, the added effort required to compile a multi-component descriptor can be avoided by utilizing a portion of a single-component curve.

Assume the curve in figure 27A is to be described mathematically and that no satisfactory matching curve has been found using $X$. As an alternative, the analyst could undertake a multiple-component description, but another single-component possibility exists.

![Figure 27](image)

Figure 27.—Using the upper half of a sigmoid: (A) the problem, and (B) a solution.

- Extend the desired curve in figure 27A such that the resulting form approximates that of a sigmoid. Add a constant to the abscissa such that $XP \pm XP$ covers the pertinent range of $X$. $(X + 10)$ turns the trick here. Also, add a suitable constant, 500 in this case, to $R$ to shift the entire sigmoid vertically to the upper-right quadrant. For representative points from the right half of this curve (for example, figure 27B), construct an overlay curve and find a matching Standard for the upper half of the overlay. The adopted single-component Standard fitted to the $X$- and $Y$-values then represents the original curve in the region $0 \leq X \leq 10$. The resulting descriptor is

$$Y = 1000e^{-x} - 500, \text{ } 0 \leq X \leq 10.$$

or

$$Y = 1000\left[1\left(\frac{(X + 10)}{20}\right) - 1\right]^{-n} \left[\frac{1}{1 - e^{-\frac{1}{n} - \frac{1}{1 - 1 - e^{-\frac{1}{n}}}}}ight].$$

- Procedures for any fraction of any form are analogous.
There may be cases where the $e^A$-transform determined from the Standards is not sufficiently accurate for the purposes of the analyst. In such cases, curve-form description procedures beyond the scope of this paper must be used.

The portion of the curve to the right of $XP$ is a mirror image of the left in $e^A$. Once $n$ and the proportional inflection point, $I$, have been selected and held constant, changes in $Y$ values depend only on departure of $X/XP$ from 1.0. Equal departures on opposite sides of 1.0 will result in equal $Y$-values. For the occasions when the analyst's curve is bell shaped and symmetrical, we really need only identify the $e^A$-transform for one side (left), as for figure 18 and in table 9 procedures.

Every $e^A$-transform is forced to zero at $X = 0$ and at $X = 2XP$ and should only be used within these limits (appendix, $e^A$ DERIVATION).

### Multiple-Component Descriptors

When a suitable match for a graphed two-dimensional hypothesis of the $X^o$-class cannot be found in the $X^o$-Standards, more satisfactory descriptors can often be developed by matching unique portions of the curve sequentially and adding descriptors for the parts to arrive at the whole. Also, development of descriptors for contiguous segments of the $X$-axis constitutes another flexible alternative.

One example, a commonly encountered kind of curve form that departs from the single-component Standards of the $X^o$-class, is represented by the graphed hypothesis in figure 28.

![Figure 28](image_url)  
*Figure 28. — A graphed hypothesis.*

Assume this is the curve to be described. The discerning analyst will immediately reject it as a single-component $X^o$ form because of the second pronounced bend in the range $700 < X < 1000$. But, let us run it through the $X^o$-single component procedures to provide a contrast for the two-component form to be developed subsequently.

<table>
<thead>
<tr>
<th>Control points</th>
<th>$X$</th>
<th>$Y$</th>
<th>Overlay control points</th>
</tr>
</thead>
<tbody>
<tr>
<td>$X$</td>
<td>$Y$</td>
<td>$X$</td>
<td>$Y$</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0.0</td>
</tr>
<tr>
<td>200</td>
<td>5.2</td>
<td>200</td>
<td>5.2</td>
</tr>
<tr>
<td>400</td>
<td>17.0</td>
<td>400</td>
<td>17.0</td>
</tr>
<tr>
<td>600</td>
<td>33.5</td>
<td>600</td>
<td>33.5</td>
</tr>
<tr>
<td>700</td>
<td>43.0</td>
<td>700</td>
<td>43.0</td>
</tr>
<tr>
<td>800</td>
<td>54.8</td>
<td>800</td>
<td>54.8</td>
</tr>
<tr>
<td>900</td>
<td>71.1</td>
<td>900</td>
<td>71.1</td>
</tr>
<tr>
<td>1000</td>
<td>100.0</td>
<td>1000</td>
<td>100.0</td>
</tr>
</tbody>
</table>

![Figure 29](image_url)  
*Figure 29. — Overlay curve for entire $X$-range compared to the $X^o$-Standards.*

There is no suitable match for the overlay curve in the $X^o$-Standards; so we resort to description of the graphed curve with two components. The first component covers the range $0 < X < 700$ since that portion of the curve in figure 28 has only one point of bend and seems reasonably similar in conformation to the $X^o$-Standards. Next, $X$ and $Y$ are scaled to 1.0 at $X = 700$ (table 11, columns 1-4).
MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES - TWO DIMENSIONS

Table 11.—Mathematical hypothesis development

<table>
<thead>
<tr>
<th>X</th>
<th>Y = a</th>
<th>Overlay control points 1st portion</th>
<th>Y'</th>
<th>Performance of Y1</th>
<th>Overlay control points 2nd portion</th>
<th>Y2</th>
<th>Performance of Y2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>0.0</td>
<td>0.00</td>
<td>0.0</td>
<td>0.00</td>
<td></td>
<td>0.0</td>
<td>0.00</td>
</tr>
<tr>
<td>200</td>
<td>5.2</td>
<td>12.0</td>
<td>5.4</td>
<td>+ 2</td>
<td></td>
<td>0</td>
<td>- 2</td>
</tr>
<tr>
<td>400</td>
<td>17.0</td>
<td>40.0</td>
<td>17.1</td>
<td>+ 1</td>
<td></td>
<td>0</td>
<td>+ 2</td>
</tr>
<tr>
<td>600</td>
<td>33.5</td>
<td>78.0</td>
<td>33.3</td>
<td>- 2</td>
<td></td>
<td>0</td>
<td>+ 2</td>
</tr>
<tr>
<td>800</td>
<td>54.8</td>
<td>100.0</td>
<td>43.0</td>
<td>0</td>
<td></td>
<td>3</td>
<td>+ 3</td>
</tr>
<tr>
<td>1000</td>
<td>71.1</td>
<td>125.0</td>
<td>53.6</td>
<td>- 1.2</td>
<td>0.80</td>
<td>0.05</td>
<td>1.4</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

As for plotting the first-portion overlay curve and comparing it to the X°-Standards (fig. 30), we find that X° gives a good match for the curve in the 0≤X≤700 range. Scaled to 43 at X = 700, X° gives a fairly close approximation of the Y1-values (table 11, column 6). As might be expected, this relatively flat curve is too flat when extended

![Figure 30. First portion overlay curve compared to the X°-Standards.](image)

![Figure 31. Second portion overlay curve compared to the X°-Standards.](image)

Mathematical descriptions of graphed hypotheses - Two Dimensions

to represent the desired curve in the 700≤X≤1000 range (column 5). The differences, 1.2, 6.0, and 22.5, represent the curve of values that must still be added to match the right side of the desired curve. Needed in the descriptor is a second component that is essentially zero in the 0≤X≤700 range and about equal to the curve of values still needed in the range 700≤X≤1000.

Plotting the overlay curve for the second curve portion and comparing it to the X°-Standards as in figure 31, we find that X° might match the required curve fairly well

![Figure 31. Second portion overlay curve compared to the X°-Standards.](image)

...but, on trial, X° performs a little better. Scaling X° to 22.5 at X = 1000 gives values close to those desired in the 700≤X≤1000 range (table 11, column 9). Then the complete descriptor is:

\[
\hat{Y} = (8.691 \times 10^{-8})X^{1.65} + (7.115 \times 10^{-5})X^{12.5},
\]

with final values listed in table 12. In this case, \( \hat{Y} \) is regarded as being satisfactorily close to \( Y \).

Table 12.—Model performance

<table>
<thead>
<tr>
<th>( X )</th>
<th>( \hat{Y} )</th>
<th>( Y )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>200</td>
<td>5.2</td>
<td>5.4</td>
</tr>
<tr>
<td>400</td>
<td>17.0</td>
<td>17.1</td>
</tr>
<tr>
<td>600</td>
<td>33.5</td>
<td>33.4</td>
</tr>
<tr>
<td>700</td>
<td>43.0</td>
<td>43.3</td>
</tr>
<tr>
<td>800</td>
<td>54.8</td>
<td>55.0</td>
</tr>
<tr>
<td>900</td>
<td>71.1</td>
<td>71.1</td>
</tr>
<tr>
<td>1000</td>
<td>100.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>
The computational ideas just shown are applicable to other curve forms, but a larger array of ideas may be necessary to independent descriptor development efforts by the reader. The examples that follow are presented with decreasing explanatory detail as seems appropriate to the stage of discussion. We will start with an abbreviated version of descriptor procedures just presented in detail for the hypothesized curve in figure 28.

**Two-Component \( \hat{X} \), No Inflection**

- Read a set of representative \( XP \) points from the graphed hypothesis (fig. 32). Select one \( X \)-value in this set, point \( d \), such that the curve over the range 0 to \( d \) appears to have about the same general conformation (one point of bend) as the \( X \)-Standards.

![Figure 32](image)

- Compare the corresponding overlay curve with the \( X \)-Standards and select an appropriate \( n \) in \( X^n \).
- Scale \( X^n \) to \( f \) at \( d, b = f/d\). Then, \( Y_1 = 2 + bX^n \). Extend \( Y_1 \) over the range \( d \) to \( e \) (dashed line).
- Compare the overlay curve for the exponential curve of residuals over the range 0 to \( e \) with the \( X \)-Standards and select an appropriate \( m \) in \( X^m \).
- Scale \( X^m \) to \( g \) at \( e, c = g/e^m \). Then, \( Y_2 = cX^m \).
- Add components \( Y_1 \) and \( Y_2 \). Then, \( Y = a + bX^n + cX^m \), and \( a = 0 \). Alternative orientations of the above curve in space can be handled simply, as specified earlier in the section entitled "TWO DIMENSIONS".

**Two-Component \( \hat{Y} \), with Inflection**

Figure 33 is a variant of the preceding sample. All steps of the descriptive process still apply, except that the second component is subtracted from the first. Then,

\[ \hat{Y} = a + bX^n - cX^m. \]

![Figure 33](image)

- Determine the residuals of the curve from the straight line for representative points over \( X \). Include three or four such points at each end of \( X \), where the curve bends away from the straight line.
- For the right-end departures, plot an overlay curve and compare to the \( X \)-Standards. Find a suitable \( n \) in \( X^n \) and scale \( X^n \) to \( h \) at \( g \), wherein \( c = h/g \). Then, \( Y_2 = cX^m \).
- For the left-end departures, plot an overlay curve and compare to the \( X \)-Standards. Reverse the \( X \)-axis by substituting \( K = (f - X) \). Plot the left-end overlay curve and compare with the \( X \)-Standards. Find a suitable \( m \) in \( X^m \) and scale \( X^m \) to \( i \) at \( j \); such that, \( d = i/j \).
- Then, the left-end descriptor is:

\[ Y_3 = dK^m. \]

Then the entire descriptor is:

\[ Y = a + bX + cX^n + dK^m, \quad 0 \leq X \leq g. \]

Variants from this curve include negative departures from either or both ends of the straight line. In such cases, the signs of the corresponding components simply become negative.

Alternative orientations of this figure in space can be handled as specified earlier in "TWO DIMENSIONS".

**Segmentation with \( X^n \)**

Where \( X^n \) does not provide a suitably accurate descriptor for \( Y \) over the entire range of \( X \), describing \( Y = f(X) \) for each of several contiguous segments of the \( X \)-range, will often remedy the problem. In figure 35:

\[ Y = 1.0 + 6.456 \times 10^{-11} (X)^3, \quad 0 \leq X \leq 60, \] results in departures from the objective curve of from 0.03 to -0.07.

\[ Y = 1.0, \quad 0 \leq X \leq 30, \] and

\[ Y = 1.0 + 5.708 \times 10^{-4} (X - 30)^2, \quad 30 \leq X \leq 60, \] results in departures of only ±0.01.
MATHEMATICAL DESCRIPTIONS OF GRAPHED HYPOTHESES - TWO DIMENSIONS

Straight-Line Slopes

For any curve that has an inflection and a distinctive peak (positive or negative), exponential and sigmoid functions may be appropriate. (Note that sigmoid functions may be created with two exponentials of opposite sign, as shown under multiple exponentials with inflection.) Sigmoid functions are shown here in combination with linear effects for added descriptive power.

In this case (fig. 36), we can use a straight line with slope \( b = \frac{f}{d} \). The intercept plus the straight line is \( Y_1 = a + bX \). Then add to \( Y_1 \) a sigmoid matching the curve differences, \( Y_d \), from the linear function, scaled to \( g \) at the point in \( X \) (e here) where \( Y_d \) peaks. (Follow the methods in table 9.) The sigmoid component is \( Y_2 = g e^{-k} \).

The final descriptor is then:

\[ \hat{Y} = a + bX + g e^{-k} \]

Alternative arrangements of the sigmoid about the straight line would require appropriate treatment (fig. 37a, b, and c).

Multiple Sigmoid

Sometimes the combination of several sigmoidal (upright or inverted) components can be used effectively to describe asymmetrical curves. Each sigmoid is matched and scaled independently and both are summed in the final descriptor.

For figure 39a, \( P = (d - X) \), for the sigmoid to the left of \( c \) and \( e' \), the sigmoid to the right of \( c \), is described in the \( X \)-scale, \( XP = d \) in the \( P \)-scale. Then,

\[ \hat{Y} = a + f e'^{-k} + g e'^{-k}, 0 \leq X \leq d. \]
For figure 39b, the descriptor is much the same as for figure 39a, but the a-value is higher and signs of the sigmoidal effects are reversed. Then,

\[ Y = \frac{\alpha - e^{-eK}}{e^{-eK} + \frac{X}{XP}}, 0 \leq X \leq d. \]

**Segmented**

A simple way of handling asymmetry is to break the X range into two parts, 0 ≤ X ≤ XP, and XP ≤ X ≤ L (fig. 40). Identify a suitable sigmoid for each part of the skewed, bell-shaped curve and scale to YP using the X-scale for the left sigmoid and the (L - X) scale for the one on the right. The result is a two-component descriptor

\[ \begin{align*}
Y_1 &= YP e^{-eK}, 0 \leq X \leq XP \\
Y_2 &= YP e^{-eK}, XP \leq X \leq L.
\end{align*} \]

**Contiguous Bell-Shaped Curves**

Bell-shaped curves, in whole or in part, can be summed in a descriptor. And, for this particular example (fig. 41), the peaks of the two leftmost bell-shaped curves occur at points other than the upper or lower extreme of the X-range, which is not true in previous examples involving sigmoids. A peculiarity of \( e^{-eK} \) is that its controlled use is within XP ± XP, where XP is the point in the X-range at which the Y-peak (either positive or negative) occurs, and XP ± XP are the points at which the tails of the bell-shaped curve drop to zero. Beyond these points, the sigmoid values become negative. Then, where an XP is located in such a way that XP ± XP does not cover the pertinent range of X, it becomes necessary to alter the X-scale so that it does. The discussion below for the foregoing graph (fig. 41) should clarify both the problem and its solution.
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - THREE DIMENSIONS

To reiterate what has been said previously, the analyst's capability for developing sensitive multidimensional hypothesis models depends largely on his or her 2-D talents. The examples shown here should provide enough procedural background to permit intuitive extension of the system to include more flexible curve-form alternatives. The live data examples included, serve to reinforce these ideas through application under real data circumstances.

Outline for "THREE DIMENSIONS"

1. Sigmoidal X1-Effect, with Constant Point of Peaking in X2 and Interacting with a Concave-Upward X2-Effect

2. Asymmetrical Bell-Shaped Ridge Over X1, with Moving Peak in X2. This is Described in Two Segments and Interacts with a Sigmoidal X2-Effect.

3. Asymmetrical Bell-Shaped Ridge Over X1, with Moving Peak in X2. This is Described in Three Segments and Interacts with a Sigmoidal X2-Effect.

4. Symmetrical, Inverse Bell-Shaped Ridge Over X1, with Constant Point of Minimum in X1 and Interacting with a Concave-Upward X2-Effect.
Concave-Upward Relation in Two Dimensions, with Constant Point of Peaking in X1 Over X2.

Sigmoidal X1-Effect, with Constant Point of Peaking in X2 and Interacting with a Concave-Upward X2-Effect

Given graphed 2-D hypotheses over X1 for representative levels of X2 (Fig. 42), and given the basic 2-D descriptor for the X1-effect (Y = intercept + scalar(X1 sigmoid)) for each curve of the set, the analyst is in a position to formulate a descriptor for the implied surface. He simply expresses the changing intercepts, scalars, and parameter(s) of the X1-sigmoid in terms of X2 and then substitutes appropriately in the basic 2-D descriptor. Pertinent information for the four sigmoid curves above has been assembled in Table 13, columns 1–6, by using information read directly from the graphs along with the X^n- and e^-X- Standards:

**Table 13.—Mathematical hypothesis development**

<table>
<thead>
<tr>
<th>X2</th>
<th>intercept</th>
<th>XP</th>
<th>n</th>
<th>l</th>
<th>scalar</th>
<th>l</th>
<th>scalar</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>20</td>
<td>10</td>
<td>3.0</td>
<td>.90</td>
<td>25</td>
<td>.90</td>
<td>25.0</td>
</tr>
<tr>
<td>2</td>
<td>20</td>
<td>10</td>
<td>3.0</td>
<td>.87</td>
<td>61</td>
<td>.87</td>
<td>61.1</td>
</tr>
<tr>
<td>4</td>
<td>20</td>
<td>10</td>
<td>3.0</td>
<td>.66</td>
<td>169</td>
<td>.66</td>
<td>169.4</td>
</tr>
<tr>
<td>6</td>
<td>20</td>
<td>10</td>
<td>3.0</td>
<td>.10</td>
<td>350</td>
<td>.10</td>
<td>350.0</td>
</tr>
</tbody>
</table>

Here, the intercept and sigmoid parameters, XP and n, are constant over X2, but the scalar increases and l decreases exponentially with increasing X2-values. Using the X^n- Standards, l = 0.9 – 0.003704(X2) and scalar = 25 + 9.028(X2)^2. These estimators perform rather well as may be seen in columns 7 and 8. Then substituting in the basic equation for Y,

\[ Y = \text{intercept} + \text{scalar}(X1 \text{ sigmoid}) \]

\[ Y = 20 + (25 + 9.028(X2)^2)(X1 \text{ sigmoid}) \]

where:

The X1 sigmoid = \[ -\frac{(X1/10 - 1)}{1 - t} \times \frac{1}{1 - t} \]

and \( t \) as specified previously, \( = 0.9 - 0.003704(X2)^2 \).

Application limits for this hypothesis are \( 0 \leq X1 \leq 10 \), and \( 0 \leq X2 \leq 6 \).

By plotting computer solutions for Y at pertinent combinations of X1 and X2, we can see that the descriptor lies quite close to the graphed curves (Fig. 42). And, the entire predicted surface appears as follows (Fig. 43):
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - THREE DIMENSIONS

Note that the original curves may be given over $X_2$ instead of $X_1$, in which case, the roles of $X_1$ and $X_2$ in descriptor development would be interchanged. Also, where multiple-component 2-D descriptors are involved, they are simply incorporated, like the single ones here, in the final descriptor.

This descriptor system is applicable to virtually any surface and is especially useful in emulating strong interactive relationships. An application to a live data set involving such interaction is documented in the next example as a means of demonstrating actual use of some of the 2-D alternatives available to the analyst.

**Sigmoidal $X_1$-Effect, with Constant Point of Peaking in $X_2$ and Interacting with a Concave-Upward $X_2$-Effect**

**Live Data Example**

Wildlife managers have expressed interest in the thermal cover provided large game animals by the live crowns of lodgepole pine trees, as reported in Cole, D. M., and Jensen 1981. Related to such cover is the height to base of live crown ($HBC$) in lodgepole shown in figure 44 where $HBC$ is expressed as a function of the average height of dominant trees ($H$) in the stand and stand density as reflected in Crown Competition Factor ($CCF$). This hypothesis reflects the expected effect for $CCF$, which is positive and sigmoidal, and that for $H$ is positive with no strong expectation as to form.

These expected effects were found to exist in a data set dedicated to the formulation of a mathematical hypothesis for the relation. The 134 observations involved, were sorted into 10-foot height groups wherein $HBC$ was plotted over $CCF$ (fig. 45).

The expected sigmoidal $CCF$-effect was clearly visible in the data and appeared to break downward from the asymptote at about the same point in $CCF$ (180-200) for all height groups. The strongest expression of the sigmoidal form appeared to exist in the

**Figure 44.** The mathematical hypothesis.

![Graphed hypothesis and model performance for partitioned data.](image)

![Graphed hypothesis and model performance for partitioned data.](image)
MATHEMATICAL DESCRIPTION OF GRAPHEO HYPOTHESES - THREE DIMENSIONS

Table 14.—Mathematical hypothesis development

\[
\begin{array}{cccccccc}
\text{H} & \text{control} & \text{Overlay} & \text{Performance} & \text{Performance} \\
\text{points} & \text{control} & \text{points} & \text{of HBC1} & \text{of HBC2} \\
\hline
n = 10, \beta = 0.183, & n = 10, \beta = 0.186, \\
XP = 500, & XP = 500, \\
a = 0, HBC1 & a = 0, HBC2 \\
\hline
\text{CCF} & \text{HBC} & \text{CCF/500} & \text{HBC18} & a + YP & \alpha^* & (HBC1 - HBC) & a + YP & \alpha^* & (HBC2 - HBC) \\
\hline
36.5 & 80 & 5.0 & 0.16 & 0.26 & 4.8 & -0.2 & 4.6 & -0.4 \\
120 & 10.1 & 0.24 & 0.56 & 11.1 & +1.0 & 10.9 & +0.8 \\
160 & 16.1 & 0.32 & 0.89 & 15.3 & -0.8 & 15.5 & -0.0 \\
500 & 18.0 & 1.00 & 1.00 & 18.0 & 0.0 & 18.0 & 0.0 \\
\hline
54.4 & 8.7 & 0.16 & 0.26 & 7.0 & -0.7 & 7.5 & -0.2 \\
120 & 16.3 & 0.24 & 0.55 & 17.4 & +1.1 & 17.5 & +1.5 \\
160 & 25.4 & 0.32 & 0.86 & 24.8 & -0.6 & 25.0 & -0.4 \\
500 & 29.5 & 1.00 & 1.00 & 29.5 & 0.0 & 29.5 & 0.0 \\
\end{array}
\]

The complete hypothesis is then:

\[
HBC = (0.12056(H)^{0.814})e
\]

with limits: 0 ≤ CCF ≤ 500; 0 ≤ H ≤ 100.

The performance of this model on the data from which it was partially derived (R^2 = 0.87) is shown as a dashed line within each H-group in figure 45. The final hypothesis presented in figure 44 is produced from model values in table 16.

Table 15.—Development of YP

<table>
<thead>
<tr>
<th>Control points</th>
<th>Overlay control points</th>
<th>b(H)^{0.814}</th>
<th>Performance of YP</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>YP</td>
<td>H80.0</td>
<td>YP48.8</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
<tr>
<td>25.1</td>
<td>9.3</td>
<td>0.31</td>
<td>0.19</td>
</tr>
<tr>
<td>36.5</td>
<td>16.4</td>
<td>0.46</td>
<td>0.24</td>
</tr>
<tr>
<td>44.6</td>
<td>21.0</td>
<td>0.56</td>
<td>0.43</td>
</tr>
<tr>
<td>54.4</td>
<td>29.4</td>
<td>0.66</td>
<td>0.60</td>
</tr>
<tr>
<td>70.2</td>
<td>41.5</td>
<td>0.86</td>
<td>0.85</td>
</tr>
<tr>
<td>74.6</td>
<td>44.5</td>
<td>0.94</td>
<td>0.91</td>
</tr>
<tr>
<td>80.0</td>
<td>48.8</td>
<td>1.00</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 16.—HBC values from the model

\[
\begin{array}{ccccccccc}
\text{CCF} & 40 & 60 & 80 & 100 & 120 & 140 & 160 & 180 & 200 & 250 & 300 & 500 \\
\hline
20 & 1 & 2 & 3 & 5 & 6 & 6 & 7 & 7 & 8 & 8 & - & - \\
30 & - & 1 & 2 & 3 & 5 & 6 & 7 & 7 & 8 & 8 & - & - \\
40 & 1 & 2 & 3 & 5 & 6 & 6 & 7 & 7 & 8 & 8 & - & - \\
50 & 1 & 3 & 7 & 12 & 16 & 20 & 23 & 25 & 26 & 27 & 27 & 27 \\
60 & 1 & 4 & 9 & 15 & 21 & 26 & 29 & 32 & 33 & 35 & 35 & 35 \\
70 & 1 & 5 & 11 & 18 & 25 & 32 & 36 & 39 & 41 & 43 & 43 & 43 \\
80 & 2 & 6 & 13 & 22 & 31 & 38 & 43 & 47 & 49 & 51 & 51 & 51 \\
90 & 2 & 7 & 15 & 26 & 36 & 45 & 51 & 55 & 57 & 60 & 60 & 60 \\
100 & 2 & 8 & 18 & 30 & 42 & 52 & 59 & 64 & 66 & 69 & 70 & 70 \\
\end{array}
\]

\[t\] Limits within which data were included.
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - THREE DIMENSIONS

The mathematical hypothesis, then, can be adjusted to data sets from different populations of trees through simple rescaling with \( b = \Sigma XY / \Sigma X^2 \), as is indicated in the TWO DIMENSIONS test.

**Symmetrical Bell-Shaped Ridge**
- Over \( X_1 \) with Moving Peak in \( X_2 \) and Interacting with a Concave-Upward \( X_2 \)-Effect

**Live Data Example**

George and Blakely (1973) reported on the calibration of fire-retardant dispersal from delivery aircraft. On each of 16 passes over a level target area, 600 gal of fire retardant were dropped from an aircraft. Square-foot coverage and volume of material reaching the ground per 100 ft\(^2\) were measured for each aerial drop along with drop height, windspeed, wind direction, aircraft speed, temperature, and humidity. Interest finally centered on the change in coverage of \( \geq 2 \) gal/100 ft\(^2\) over a controlled range of drop heights within a partially controlled range of windspeeds. (The effects of all other variables were either negligible or unidentifiable in this data set).

If the plane were rolling along the ground (drop height = 3 ft) at normal flying speed and in the absence of wind, it would be expected that a 600-gal drop would be distributed over a relatively narrow strip of ground and that the area covered by \( \geq 2 \) gal/100 ft\(^2\) would be held to some nominal value. Increased aircraft (and drop) height should result in greater dispersal; maximum coverage should be reached at some optimal height. As drop height is increased beyond the optimum, dispersal of the retardant from air friction and evaporation should become more complete; coverage should finally reach zero at some relatively great drop height.

So expectation at zero wind is for a bell-shaped curve truncated to the left of the peak. With increasing wind, optimal drop height and peak coverage should diminish, since dispersal of the retardant would be accentuated by the horizontal shearing force of the wind. Thus, the expected surface would be as shown in figure 47.

**Figure 47.** The general expected relation.

**Figure 48.** Expected trends fitted through the data.

**Figure 49.** The graphed hypothesis.
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - THREE DIMENSIONS

Figure 49, then, is the basic hypothesis for which a descriptor was developed. Assuming that the curves over $H$ (fig. 48) could be suitably represented by segments of symmetrical bell-shaped forms, $e^{-x}$ was deemed applicable and the identification of matching e-transforms was begun.

The e-transform is limited in that it will yield values of from zero to one in the full bell-shaped form only within the range $HP < HP$, where $HP$ is the point in $H$ at which the $C$-values peak. And, it was a unique feature of this example that the maximum $H$-values for all three curves of the set exceed $2HP$ (fig. 50).

![Graphed trends from figures 49 and 50, extended.](image)

The broadest curve (for $W = 0.5$) was estimated to peak at $H = 273$. In this case, the matching $e$-transform would have dropped below zero where $H$ exceeded $2HP$, or $2(273) = 546$, long before the approximate graphed limit of $H = 1150$ was reached. One solution here is to reverse the $H$-scale using as a base, the largest $H$-value of the broadest sigmoid likely to be encountered in the set of interest. Since an even broader sigmoid than $W = 0.5$ is expected to exist at $W = 0$, the largest $H$-value is estimated at $1150 + 100 = 1250$. The $H$-axis is then reversed using $X_1 = (1250 - H)$ where, as may be seen in figure 50, $XPI = 977$, 1072, and 1170 for $W = 0.5$, $W = 3.2$, and $W = 9.6$, respectively. Now, as required for each sigmoid, $XPI \pm XPI$ includes both the largest and smallest value of $X_1$ that will be encountered, and $X_1 = (1250 - X)$ is used as the basis for the bell-shaped $H$-effect descriptor ($e^{-x}$). The coverage estimator is then specified as,

$$C = \frac{\left(\frac{X_1}{XPI}\right) - 1}{1 - I} - e^{-(1/(1-I))^n} - e^{-(1/(1-I))^n}$$

where,

$C$ = coverage in square feet ($\geq 2$ gal/100 ft)

$X_1 = (1250-H)$

$XPI$ = the point in $X_1$ at which $C$ peaks

$CPI$ = coverage peak at $XPI$, scalar for the sigmoid

$I$ = the proportional point in $X_1$ at which the inflection point of the bell-shaped curve occurs (a) in the range $X_1 = 0$ to $XPI$ for the left half of the curve or (b) in the range $XPI$ to $2XPI$ for the right half (fig. 50).

$n$ = the power of the transform that dictates the degree of curvature above and below any inflection point.

Since the $XPI$- and $CPI$-values read from the three curves in figure 50 varied with wind in accord with expectation, they were each expressed as a suitable function of wind.

Multiple-component $X^*$-descriptors were utilized. $XPI$- and $CPI$-values from the foregoing equations along with representative paired $X$- and $C$-values read from the left side of each bell-shaped curve were the basis for overlay curves. Standards for $e^{-x}$ that were suitably similar to the overlay curves were identified, and the corresponding $n$ (field constant) and $I$-values were recorded. Also, $y$, varying over wind in accord with expectation, was described as a function of wind to complete the inputs to the coverage estimator. The details of these computations follow.

For $XPI = f(W)$

$XPI$ and $XPI$-values (0.5, 977; 3.2, 1072; and 9.6, 1170) were read from figure 50. These points were plotted and a smooth curve was drawn through them (fig. 51) that extended over the relevant range of $W$ ($0 - 15$ mi/h).

The initial descriptor approach here was to subtract a single-component exponential curve on the reversed $W$-scale from the intercept, 1190 (table 17, columns 1-7). $W$ was reversed in $X_1 = (15 - W)$ to align large values of the independent variable with those of $XPI$. Next, the three data point $XPI$-values and associated coverage differences (abscissa) from 1190 were scaled to 1.0 at $X_1 = 14.5$. The overlay curve was not well matched by any curve in the $X^*$-Standards; so a two-component approach was adopted.
Table 17.—Descriptor development for XP1

<table>
<thead>
<tr>
<th>W</th>
<th>XPI</th>
<th>(15 - W)</th>
<th>(1190 - XPI)</th>
<th>Overlay control points</th>
<th>( b(XPI)^{25} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>15.0</td>
<td>15.0</td>
<td>15.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>0.5</td>
<td>1977</td>
<td>14.5</td>
<td>118</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>3.2</td>
<td>1170</td>
<td>11.6</td>
<td>18</td>
<td>0.46</td>
<td>0.46</td>
</tr>
<tr>
<td>6.7</td>
<td>1170</td>
<td>15.4</td>
<td>20</td>
<td>0.17</td>
<td>0.17</td>
</tr>
<tr>
<td>15.0</td>
<td>1170</td>
<td>9.6</td>
<td>0.0</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>X</th>
<th>Y</th>
<th>X1</th>
<th>Y1</th>
<th>X1/XP1</th>
<th>Y1/YP1</th>
<th>Y1</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
<td>(6)</td>
<td>(7)</td>
</tr>
<tr>
<td>15.0</td>
<td>1170</td>
<td>9.6</td>
<td>0.0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.0</td>
</tr>
<tr>
<td>11.5</td>
<td>1977</td>
<td>14.5</td>
<td>25</td>
<td>1.00</td>
<td>1.00</td>
<td>25</td>
</tr>
<tr>
<td>11.8</td>
<td>1170</td>
<td>11.6</td>
<td>0</td>
<td>0.81</td>
<td>0.81</td>
<td>0</td>
</tr>
<tr>
<td>5.4</td>
<td>1170</td>
<td>9.6</td>
<td>0</td>
<td>0.37</td>
<td>0.37</td>
<td>0</td>
</tr>
<tr>
<td>0.0</td>
<td>1170</td>
<td>9.6</td>
<td>0</td>
<td>0.00</td>
<td>0.00</td>
<td>0.0</td>
</tr>
</tbody>
</table>

The completed function was then

\[ \hat{X}_{PI} = 1190 - 0.4572(15 - W)^{25} - 1.4811 \times 10^{-22}(15 - W)^{25}, 0 \leq W \leq 15. \]

For \( CP1 = f(W) \)

Paired \( W \) and \( CP1 \) values (0.5, 8900; 3.2, 8100; and 9.6, 6970) were read from figure 50. These points were plotted and a smooth curve was drawn through them (fig. 53) that extended over the relevant range of wind, 0-15 m/s.

Figure 52.—First- and second-component overlay curves compared to the \( X^* \) -Standards.

Figure 53.—A flat curve, \( X^1 \), matches the lower end (9.6 = \( W \leq 15.0 \)) of the graphed hypothesis.

The completed function was then

\[ \hat{X}_{PI} = 1190 - 0.4572(15 - W)^{25} - 1.4811 \times 10^{-22}(15 - W)^{25}, 0 \leq W \leq 15. \]
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES – THREE DIMENSIONS

Again, a single-component exponential was first tried for the descriptor here (table 18, columns 1-7). Reversing the W-axis to $X = (15 - W)$ to align large values of the independent variable with those of CP1, the X1- and CP1-values were scaled to 1.0 at $X1 = 14.5$. An overlay curve was constructed and compared to the $X^2$-Standards. But, none of the Standards were acceptably close to the overlay, so a two-component model was tried next (table 18).

To start with, the lower end of the curve ($6.0 \leq X1 \leq 15.4$) was fitted with an intuitively selected flat form ($X1^2$); it was obvious without scaling and overlays that the curve was extremely flat in that range of $X1$ (fig. 53). The component $X1^{1.5}$ was then scaled to the difference (580) between the intercept ($6390$) and the desired curve height ($6970$) at $X1 = 5.4$. So this first component (table 18, column 7), along with the intercept, $6390$, constitutes the partial descriptor:

$$\hat{CP1} = 6390 + b(X1)^{1.5}$$

Table 18.—Descriptor development for CP1

<table>
<thead>
<tr>
<th>$W$</th>
<th>CP1</th>
<th>$X1$</th>
<th>Y1</th>
<th>$X1^{1.5}$</th>
<th>Y1$^{1.5}$</th>
<th>Y1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>6900</td>
<td>14.5</td>
<td>2510</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>5</td>
<td>8100</td>
<td>11.8</td>
<td>1710</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>9.6</td>
<td>6970</td>
<td>5.4</td>
<td>586</td>
<td>1.00</td>
<td>1.00</td>
<td>586</td>
</tr>
<tr>
<td>15.0</td>
<td>6390</td>
<td>0.0</td>
<td>0.00</td>
<td>--</td>
<td>--</td>
<td>--</td>
</tr>
</tbody>
</table>

The curve $n = 4.0$ matched the overlay fairly well. So ($X1^2$) was scaled to the difference, 791, at $X1 = 14.5$ (table 18, column 12). This was added to the intercept and the first component to arrive at the two-component descriptor, $CP1 = 6390 + 90.739(15 - W)^{1.5} + 0.01794(15 - W)^2$, $6 \leq W \leq 15$. Note that the related estimates in table 18, column 13 were regarded as being satisfactorily close to the original objective values in column 2.

Matching the Bell-Shaped Curves

Tables 19 through 21 list XPI- and CP1-values from the equations developed in the foregoing sections along with some representative $H$- and $C$-values from the data-based portions of the bell-shaped curves in figure 50. Also listed are $C$-value estimates obtained by using matching $e^{-k}$-Standards (see the overlay curves in figure 55).

As may be seen in tables 19-21, the coverage values of the curve for each wind group were scaled to 1.0 by using the corresponding CP1 as the maximum. The values of $X1$ of $(2XPI - X1)$ for each wind group were similarly scaled by using the corresponding XPI as the maximum. After the paired proportions for $H$-levels within each wind group were plotted and smoothed, the resulting overlay curves were compared to the $e^{-k}$-Standards for $n = 2$ (fig. 55). Although better matches for some of these overlays can be found in Standards that have $n$ other than 2.0, it was decided to hold $n$ constant.
### Table 19 — Descriptor development for C, W = 0.5

<table>
<thead>
<tr>
<th>W</th>
<th>H</th>
<th>C</th>
<th>X1</th>
<th>T</th>
<th>Y1</th>
<th>Y1/8900</th>
<th>Y1</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>50</td>
<td>6680</td>
<td>1200</td>
<td>754</td>
<td>6680</td>
<td>0.77</td>
<td>0.77</td>
</tr>
<tr>
<td>100</td>
<td>7860</td>
<td>1150</td>
<td>864</td>
<td>7860</td>
<td>0.82</td>
<td>0.88</td>
<td>7793</td>
</tr>
<tr>
<td>150</td>
<td>8450</td>
<td>1100</td>
<td>854</td>
<td>8450</td>
<td>0.87</td>
<td>0.95</td>
<td>8322</td>
</tr>
<tr>
<td>200</td>
<td>8770</td>
<td>1056</td>
<td>964</td>
<td>8770</td>
<td>0.93</td>
<td>0.99</td>
<td>8692</td>
</tr>
<tr>
<td>273</td>
<td>8260</td>
<td>917</td>
<td>957</td>
<td>8260</td>
<td>1.00</td>
<td>1.00</td>
<td>8900</td>
</tr>
</tbody>
</table>

1. Temporary transform for plotting the overlay curve. The temporary X-values (T) must be in ascending order toward XPI/ = 977 to match the position of the e =. Standards in two-space.
2. All symbols are left- or right-handed, as (C - a) = n = 71.
3. Estimated coverage using e = with X1 = (1250 - H), f = 0.51, n = 2.0, XPI = 977, CPI = 8900, and a zero intercept.
4. Smoother values from XPI and CPI, each a function of wind (shown earlier in the text).

#### Table 20. — Descriptor development for C, W = 3.2

<table>
<thead>
<tr>
<th>W</th>
<th>H</th>
<th>C</th>
<th>X1</th>
<th>T</th>
<th>Y1</th>
<th>T/072</th>
<th>Y1/8107</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.2</td>
<td>50</td>
<td>6730</td>
<td>1200</td>
<td>944</td>
<td>6730</td>
<td>0.88</td>
<td>0.83</td>
</tr>
<tr>
<td>100</td>
<td>7860</td>
<td>1150</td>
<td>944</td>
<td>7860</td>
<td>0.93</td>
<td>0.94</td>
<td>7612</td>
</tr>
<tr>
<td>150</td>
<td>8450</td>
<td>1100</td>
<td>1044</td>
<td>8450</td>
<td>0.97</td>
<td>0.99</td>
<td>8042</td>
</tr>
<tr>
<td>176</td>
<td>8260</td>
<td>917</td>
<td>957</td>
<td>8260</td>
<td>1.00</td>
<td>1.00</td>
<td>8107</td>
</tr>
</tbody>
</table>

1. Estimated coverage using e = X1 with X1 = (1250 - H), f = 0.71, n = 2.0, XPI = 1072, CPI = 8107, and a zero intercept.

### Table 21. — Descriptor development for C, W = 9.6

<table>
<thead>
<tr>
<th>W</th>
<th>H</th>
<th>C</th>
<th>X1</th>
<th>Y1</th>
<th>X1/1170</th>
<th>Y1/6895</th>
</tr>
</thead>
<tbody>
<tr>
<td>9.6</td>
<td>80</td>
<td>6985</td>
<td>1170</td>
<td>6985</td>
<td>1.00</td>
<td>1.00</td>
</tr>
<tr>
<td>100</td>
<td>6850</td>
<td>1100</td>
<td>6850</td>
<td>0.98</td>
<td>0.98</td>
<td></td>
</tr>
<tr>
<td>150</td>
<td>5300</td>
<td>1100</td>
<td>5300</td>
<td>0.94</td>
<td>0.76</td>
<td></td>
</tr>
<tr>
<td>200</td>
<td>3270</td>
<td>1050</td>
<td>3270</td>
<td>0.90</td>
<td>0.47</td>
<td></td>
</tr>
<tr>
<td>250</td>
<td>1830</td>
<td>1000</td>
<td>1830</td>
<td>0.85</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td>300</td>
<td>997</td>
<td>950</td>
<td>997</td>
<td>0.81</td>
<td>0.14</td>
<td></td>
</tr>
</tbody>
</table>

1. Since the right half of this curve (Fig. 50) was read, X1-values are already ascending toward XPI = 1170 and no temporary transform of X1 is necessary for plotting the overlay curve.
2. Estimated coverage using e /XP = 1250 - H, f = 0.85, n = 2.0, XPI = 1170, CPI = 6895, and a zero intercept.

---

**Figure 55.** Standards for n = 2.0 and proportional departure of inflection points in X, (X1), from XP, (X1/XP).

---

over wind. Then, the I-values corresponding to the best matching curve alternatives in this set, were estimated to be 0.51, 0.71, and 0.88 for W = 0.5, 3.2, and 9.6, respectively. On scaling and checking these alternative curves (by means of a small computer) at the tabled H-values within W-group, some improvement was achieved with the final I-array, 0.51, 0.71, and 0.88. The last column in each of tables 19-21 shows the final scaled coverage estimates. There was close proximity of these values to actual coverage, C, so I = f(W) was developed next.
For $I = f(W)$

Using the paired wind and $I$-values just determined (0.5, 0.51; 3.2, 0.71; and 9.6, 0.88), $I$ was plotted over $W$ and a smooth initial curve drawn through the three data points (fig. 56).

On checking the performance of $e^{-N}$ over the range of $W$, with smoothed CP-, XP-, and initial $I$-values, it was found that the left edge of the surface undulated unacceptably (at $H=50$, $0 \leq W \leq 5$). $I$-values were adjusted iteratively to minimize the problem and the new (final) set of points shown in figure 56 were established. The smooth (final) curve through these points became the objective curve for which a descriptor was developed.

This curve was judged to require a multicomponent model because neither the sigmoid nor the exponentials have the flattened central segment found there. The natural segment was represented by a straight line; the negative differences at the left end by an exponential; and those at the right by the lower portion of a sigmoid (fig. 57).

The Flat Segment.—A straight line was drawn along the final curve. The intercept, $a = 0.62$, and values necessary to the coefficient for $W$, $(1.054 - 0.620)/15 = 0.0289$, were read directly from the graph. Then $Y_1$, the first descriptor component, $a = 0.620 + 0.0289(W)$, (see table 22, column 3).

The Left End.—This residual portion of the curve ($Y_2$) was satisfactorily matched by a single exponential of the reversed $W$-axis ($Y_2$), scaled at $X_1 = 14.5$ (table 22, columns 4–8).

So the partial form at this point is:

$$Y_2 = 0.62 + 0.0289(W) - 4.3703 \times 10^{-21}(15 - W)^{15.75}.$$ 

The Right End.—For this portion of the curve, the differences ($Y_3$) from the straight line and associated scaling information are listed in table 23. It is apparent in figure 58 (right side) that the objective curve is not well matched by any simple power of $W$. The same held true for a multiple-component descriptor that was subsequently developed. By

Table 22.—Descriptor for $I$, the left end and central segment

<table>
<thead>
<tr>
<th>$W$</th>
<th>$Y_1$</th>
<th>$Y_2$</th>
<th>$X_1$</th>
<th>$Y_3$</th>
<th>$X_2$</th>
<th>$Y_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>0.620</td>
<td>0.620</td>
<td>15.0</td>
<td>0.220</td>
<td>1.00</td>
<td>0.01</td>
</tr>
<tr>
<td>2.0</td>
<td>0.621</td>
<td>0.624</td>
<td>14.0</td>
<td>0.061</td>
<td>0.97</td>
<td>0.05</td>
</tr>
<tr>
<td>3.2</td>
<td>0.687</td>
<td>0.678</td>
<td>13.0</td>
<td>0.021</td>
<td>0.90</td>
<td>0.17</td>
</tr>
<tr>
<td>3.2</td>
<td>0.710</td>
<td>0.712</td>
<td>11.8</td>
<td>0.002</td>
<td>0.81</td>
<td>0.02</td>
</tr>
</tbody>
</table>

1 Although not shown in the text here, the curve implied by the overlay points matched 16:1:17 in the $x^2$.

Standards very well. Iterative checking showed $n = 18.75$ to perform best within that range. Then, $e = 0.124(14.5)^{15.75}$ $= 4.3703 \times 10^{-21}$. 

The Right End.—For this portion of the curve, the differences ($Y_3$) from the straight line and associated scaling information are listed in table 23. It is apparent in figure 58 (right side) that the objective curve is not well matched by any simple power of $W$. The same held true for a multiple-component descriptor that was subsequently developed. By
reorienting the exponential curve in space, however, reversing $W$ to $Xl = (15 - W)$ and inverting $Y3$ to $Y4 = (0.15 - Y3)$, the curve becomes convex-upward with increasing values of $X1$. Also, it is at least roughly matched by the sigmoid Standard, $n = 8$ and $l = 0.108$ (see table 24 and the left overlay curve in fig. 59). This sigmoid was the result of an interpolation between the two Standards $l = 0.1$ and 0.2 at $n = 8$.

Then, the right end was specified as:

$$Y4 = (0.15 - Y3) = 0.150(e^{-k})$$

$$Y3 = 0.150 - 0.150(e^{-k})$$

and simplified:

$$Y3 = 0.150 - 0.150 \left[ \frac{15}{15 - 0.108} - \frac{1}{1 - 0.108} \right]$$

The overall performance of this descriptor can be seen from the tabulation of the objective and estimated $l$ in table 25, where $f$ and $l$ were suitably close over the range of $W$, as already verified by segment during development.
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - THREE DIMENSIONS

Table 25. - The performance of i

<table>
<thead>
<tr>
<th>W</th>
<th>Objective</th>
<th>i</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.0</td>
<td>0.400</td>
<td>0.401</td>
</tr>
<tr>
<td>5</td>
<td>0.510</td>
<td>0.510</td>
</tr>
<tr>
<td>1.0</td>
<td>0.581</td>
<td>0.580</td>
</tr>
<tr>
<td>2.0</td>
<td>0.657</td>
<td>0.658</td>
</tr>
<tr>
<td>3.2</td>
<td>0.710</td>
<td>0.709</td>
</tr>
<tr>
<td>5.0</td>
<td>0.766</td>
<td>0.764</td>
</tr>
<tr>
<td>7.5</td>
<td>0.838</td>
<td>0.835</td>
</tr>
<tr>
<td>8.4</td>
<td>0.863</td>
<td>0.859</td>
</tr>
<tr>
<td>9.6</td>
<td>0.886</td>
<td>0.889</td>
</tr>
<tr>
<td>10.4</td>
<td>0.904</td>
<td>0.900</td>
</tr>
<tr>
<td>13.2</td>
<td>0.904</td>
<td>0.905</td>
</tr>
<tr>
<td>15.0</td>
<td>0.904</td>
<td>0.904</td>
</tr>
</tbody>
</table>

Summary for C:
At this point, all necessary inputs have been derived and the complete hypothesis is:

\[
C' = CPI \left(1 - \frac{(X/X') - 1}{1 - \sigma^2} \right)^2
\]

where:

- \( C' \) = coverage in square feet, estimated from the preleast squares fit model
- \( CPI \) = coverage peak for any specified windspeed, \( W \), and
- \( CPI = 6390 + 90.739(15 - W)^2 + 0.017894(15 - W)^3, \) for \( W \leq 15 \)
- \( X1 = (1250 - H) \)
- \( H \) = drop height, \( 50 \leq H \leq 300 \)
- \( XPI = \) point in \( X1 \) at which \( CPI \) occurs, or
- \( XPI = 1190 - 0.4572(15 - W)^2 - 1.4411 \times 10^{-8}(15 - W)^3 \)
- \( I = \) inflection point in \( X \) expressed as a proportion of \( XPI, \) \( (X1/XP) \) in Matcha-curve-1. Jensen and Homeyer 1970,
- \( 0.62 + 0.0289(W) - 4.3703 \times 10^{-13}(15 - W)^{-1} \)
- \( \sigma = 0.892 \)

The hypothesis was then adjusted back to the data from which it was partially derived following the fitting procedures discussed under 2-D, \( X^2 \) and \( e = \sigma \), where \( C' = X \) and the simple least squares adjustment coefficient for the entire hypothesis,

\[
b = \Sigma XY \Sigma X^2 = 1.000569 \text{ and } C = bC'.
\]

Figure 55. - Hypothesis performance by wind group.

Recall that through the use of controlled \( i \)-values (fig. 56) for curves passing close to the three high-wind points, low variation of these points from the surface was assured. Alternative \( i \)-values could be used that would allow for variation from the surface, in accord with some assumed variance criteria. This refinement was not attempted and the model was adopted as developed. Predicted values (table 26) and the associated surface (fig. 60) for the model are as shown below.

Table 26. - Model values

<table>
<thead>
<tr>
<th>Drop height</th>
<th>Wind</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>50</td>
</tr>
<tr>
<td>5</td>
<td>6.830</td>
</tr>
<tr>
<td>10</td>
<td>6.617</td>
</tr>
<tr>
<td>15</td>
<td>6.345</td>
</tr>
</tbody>
</table>
Asymmetrical Bell-Shaped Ridge over X1, with Moving Peak in X2. This is Described in Two Segments and Interacts with a Sigmoidal X2-Effect

In this hypothesis from Klein, Parker, and Jensen 1978, tree mortality percent in a western forest is characterized by tree diameter (d.b.h.) over the course of a beetle epidemic (fig. 61).

The graphed hypothesis is applicable only at discrete points in time and at the midpoints of 2-inch d.b.h. classes, but 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

It is appropriate to review the imposing list of constraints that can be met in models like this one. Controllable items include: the intercept for the whole surface; the elevation, breadth, and positioning of the peaks over the interacting independent variables; the trend in coverage at the truncated left edge of the surface; the magnitude of curvature above and below the inflection points in the bell-shaped curves; and the positioning of the inflection points themselves within the range of the independent variables.

**Live Data Example**

In the 3-D hypothesis, the X, Y, and Z variables are given by:

\[
Z = f(X, Y) \]

where \( Z \) is the dependent variable, \( X \) and \( Y \) are independent variables. The function \( f \) represents the mathematical model describing the relationship between the variables.

The graphed hypothesis is applicable only at discrete points in time and at the midpoints of 2-inch d.b.h. classes, but 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 the left-end intercept, \( Int \), or floor, upon which the left-side sigmoidal curves rest conceptually. \( Int \) is a function of d.b.h., as is \( XP \), the location of the ridge in time.
MATHEMATICAL DESCRIPTION OF GRAPHAED HYPOTHESES - THREE DIMENSIONS

Sigmoid curves are described within the range $XP \pm 5$ years ($Yr$), since all sigmoids are estimated to be fully expressed therein. Use of a constant maximum range, such as $XP \pm 5$ 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

$Lsig = \text{left sigmoid} = f_d(Ln, Ll, XP, Yr)$. See the basic sigmoid parameters defined in Jensen 1979.

$Ln = \text{Lsig power} = f_{d} (d.b.h.)$
$Ll = \text{Lsig inflection point as a proportion of the range in years from} \, XP - 5 \, \text{to} \, XP$
$= f_{d}(d.b.h.)$

$XP = \text{point in time where} \, p_x = f_d(d.b.h.)$

$Rsig = \text{right sigmoid} = f_d(Rn, Rl, XP, Yr)$
$Rn = \text{Rsig power} = f_{d}(d.b.h.)$
$Rl = \text{Rsig inflection point as a proportion of the range in years from} \, XP + 5 \, \text{to} \, XP$
$= f_{d}(d.b.h.)$

$Int = \text{intercept, left edge} = f_{d}(d.b.h.)$
$Rsc = \text{Rsig scalar} = \text{ridge top} = f_{d}(d.b.h.)$
$Lsc = \text{Lsig scalar} = (\text{ridge bottom}) - (\text{Int})$

Segmental constraints:

Left side: $66 \leq Yr \leq XP$, discrete values only

Right side: $XP \leq Yr \leq 71$, discrete values only

Either side: $8 \leq \text{d.b.h.} \leq 18$, midpoints of 2-inch d.b.h. class 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.

The explicit hypothesis and moderate developmental detail follow:

Mortality percent $= K[Int + Lsc(Lsig) + Rsc(Rsig)]$

where

$K = \text{least-squares coefficient} = 0.9877$

$Int = 1 + 0.2321(\text{d.b.h.} - 8)^3^{1.8}$

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

$Lsc = \text{Rsc} - \text{Int}$

$Lsig = e^{-\frac{5 - [Yr - XP]}{1}} + \frac{\frac{1}{1 - LI}}{[1 - e^{-\frac{5}{1 - LI}}]^n} + \frac{\frac{1}{1 - LI}}{[1 - e^{-\frac{5}{1 - LI}}]^n}$

$XP = 67.65 + 1.2257 e^{-\frac{20 - (\text{d.b.h.})}{12} - \frac{1}{0.65}}$

$Ll = \frac{0.851 - 0.1531 e^{-\frac{7 - (\text{d.b.h.}) - 13.5}{7}}}{0.45}$

$Ln = 1.5$

$Rsig = \text{as for} \, Lsig$, but using $Rl$ and $Rn$

$Rl = 0.752 + 7.8173 \times 10^{-4} (\text{d.b.h.} - 14)^{2.8}$

$Rn = 2$

Limits

$66 \leq Yr \leq 71$, discrete units only

$66 \leq Yr \leq XP$, for Int, Lsc, Lsig

$XP < Yr \leq 71$, for Rsc, Rsig

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

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. 66). This was satisfactorily described by using the 4th Standards.

$Rsc$ is simply the height of the ridge above zero (no intercept) and serves as the scalar for the right-hand 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. 65). Note that the ridge line in figure 62 only appears to be sigmoidal by reason of the sigmoidal change of point-of-peaking in time with a change in diameter.

$Lsc = \text{Rsc} - \text{Int}$ and is the scalar for the left-hand 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. 66). This curve was estimated to asymptote at 68.88 for lower diameters and 67.75 for larger diameters. A suitable match was

---

MATHEMATICAL DESCRIPTION OF GRAPHAED HYPOTHESES - THREE DIMENSIONS

Supplementary Explanatory Detail

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.

The explicit hypothesis and moderate developmental detail follow:

Mortality percent $= K[Int + Lsc(Lsig) + Rsc(Rsig)]$

where

$K = \text{least-squares coefficient} = 0.9877$

$Int = 1 + 0.2321(\text{d.b.h.} - 8)^3^{1.8}$

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

$Lsc = \text{Rsc} - \text{Int}$

$Lsig = e^{-\frac{5 - [Yr - XP]}{1}} + \frac{\frac{1}{1 - LI}}{[1 - e^{-\frac{5}{1 - LI}}]^n} + \frac{\frac{1}{1 - LI}}{[1 - e^{-\frac{5}{1 - LI}}]^n}$

$XP = 67.65 + 1.2257 e^{-\frac{20 - (\text{d.b.h.})}{12} - \frac{1}{0.65}}$

$Ll = \frac{0.851 - 0.1531 e^{-\frac{7 - (\text{d.b.h.}) - 13.5}{7}}}{0.45}$

$Ln = 1.5$

$Rsig = \text{as for} \, Lsig$, but using $Rl$ and $Rn$

$Rl = 0.752 + 7.8173 \times 10^{-4} (\text{d.b.h.} - 14)^{2.8}$

$Rn = 2$

Limits

$66 \leq Yr \leq 71$, discrete units only

$66 \leq Yr \leq XP$, for Int, Lsc, Lsig

$XP < Yr \leq 71$, for Rsc, Rsig

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

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. 66). This was satisfactorily described by using the 4th Standards.

$Rsc$ is simply the height of the ridge above zero (no intercept) and serves as the scalar for the right-hand 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. 65). Note that the ridge line in figure 62 only appears to be sigmoidal by reason of the sigmoidal change of point-of-peaking in time with a change in diameter.

$Lsc = \text{Rsc} - \text{Int}$ and is the scalar for the left-hand 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. 66). This curve was estimated to asymptote at 68.88 for lower diameters and 67.75 for larger diameters. A suitable match was
MATHEMATICAL DESCRIPTION OF GRAPHERD HYPOTHESES – THREE DIMENSIONS

found in the $e^{-x}$-Standards when the d.b.h. scale was reversed to 20 (d.b.h.).

$L_{sig}$ specifies the sigmoidal shape of the left side (fig. 62). 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 were plotted for these cross sections (fig. 67). The pattern was one of wider sigmoidal crowns near d.b.h. = 14 inches, narrowing with departure from that d.b.h. This is reflected in the bell-shaped function for the inflection points, $L_{I}$. Also, the $e^{-x}$-Standards with $n = 1.5$ were found to represent this curve array with reasonable accuracy, so $L_{n}$ was set at the constant, 1.5.

$L_{I}$, the function for the inflection points, is the dashed line in figure 68 and represents the inflection points adopted and plotted for each overlay curve. $L_{I}$ reflects the width-of-curve-crown trend noted under $L_{sig}$ and in the text.
Asymmetrical
Bell-Shaped Ridge
over X1, with
Moving Peak in X2.
This is described in
Three Segments and
Interacts with a
Sigmoidal X2-Effect

Quality score \( t \) was originally described as an aggregate of 11 planar regions over
flow \( F \) and stability \( S \) of gap-graded road materials, a strongly segmented descriptor
(fig. 69, from Lee and others 1973). The objective here was to smooth the figure mathematically. Although the problem is an unusual one, its solution serves admirably to
demonstrate descriptor segmentation.

From figure 69, it can be seen that opposite sides of the ridge differ substantially in
slope so that a symmetrical, flat-topped, bell-shaped curve would be required to
describe the cross section at any point in \( S \) while rounding the junctions of planes. A rela-
tively simple descriptor alternative involves segmentation of the \( S \) and \( F \) 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 appro-
priate). 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 69. — Quality of gap-graded road materials: segmented
planar form.

Figure 70. — Quality of gap-graded road materials: segmented
planar and final descriptor forms.
MATHEMATICAL DESCRIPTION OF GRAPHS HYPOTHESES - THREE DIMENSIONS

Explicit hypothesis and moderate developmental detail follow:

If

\[ F > LO, \quad I = \text{Trig}(LO) \]
\[ LO \leq F \leq RO, \quad I = 10 \cdot \text{Trig} \]
\[ F < RO, \quad I = \text{Trig}(RO) \]

where

\[ F = \text{soil flow} \]
\[ LO = 8.988 + 0.0017036(S) \]
\[ RO = 6.404 + 0.0004876(S) \]
\[ S = \text{soil stability} \]

\[ \text{Trig} = 10e^{ \frac{(LO + 20 - F) - 1}{20} } \]
\[ 0.23 \]
\[ \text{Trig} = 1 - e^{ \frac{(6000 - S) - 1}{6000} } \]
\[ 0.185 \]
\[ \text{Trig} = 1 - e^{ \frac{[RO - 11] + F - 1}{11} } \]
\[ 0.30 \]

Limits

\[ 0 \leq F \leq 28, \]
\[ 0 \leq S \leq 5000. \]

Supplementary Explanatory Detail

\[ \text{Trig}, \] the left segment of figure 69, has a constant cross section over \( F \) for \( 1240 \leq S \leq 5000, \) approximated in the descriptor by a sigmoid oriented at \( F = LO \) and extending leftward a distance of 20 units in \( F. \) This is about the minimum operational span for the sigmoid here. At the upper extreme of \( S, (S = 5000), \) the left-segment sigmoid is expected to be completely specified in the range \( 27 \geq F \geq LO, \) as in figure 70. Setting 28 as the upper limit 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 71, to associate the largest value of the sigmoidal span, 20, with the peak of the sigmoid at \( LO. \) By scaling control points from the left cross section at \( S = 1240 \) (fig. 71) to 1.0 in \( X. (LO + 20 - F), \) and \( Y, \) (score), and making an overlay curve, an appropriate sigmoid was identified from the e-Standard.

Summary

The final surface, then, is simply the sum of the three contiguous ridge segments, all of which are truncated at appropriate points in \( S \) through multiplication by the proportional values of \( \text{Trig}. \)
MATHEMATICAL DESCRIPTION OF GRAPHE D HYPOTHESES - THREE DIMENSIONS

Symmetrical, Inverse Bell-Shaped Ridge over X1, with Constant Point of Minimum in X1 and Interacting with a Concave-Upward X2-Effect

Live Data Example

Here, the amount of natural reproduction (number of lodgepole pine seedlings per acre) over the Sleeping Child burn was expected to be affected by both azimuth and slope (data from Lyon 1976). Further, azimuth and slope were expected to interact strongly, the magnitude of response to azimuth being dependent on slope percent.

A small data set was used to develop a graphic hypothesis consistent with these expectations. It was very close to the mathematical hypothesis pictured in figure 73.

The Azimuth (A)-effect may be seen as cross sections over AZ at slope percent (SL) = 20, 30, 35, and 35. These were screened for shape using the $e^{-A}$-Standards. Note that, since the bell-shaped curves were symmetrical, either sigmoidal half would be representative of the whole (the left-hand was used). Also, negative departures from either edge could be used to specify the depth of the sigmoids (these negative values were ignored in the screening, but were accounted for in the final mathematical hypothesis). Sigmoidal were found to be sufficiently similar at different SL-levels to adopt the common form,

$$e^{-A} = 1.0675e^{\frac{AZ}{180}} - 0.655$$

where the trough is at a fixed point in $AZ$ ($XP = 180$), here.

The trough of the objective figure (MIN) was expressed as the height of the negative bell-shaped (or sigmoidal) peaks above the 3-D base. As with $EDGE$, this was described as the lower 85 percent of a sigmoid on the reversed slope scale with a maximum value of 100. And,

$$MIN = \frac{2500}{1000}e^{\frac{(100 - SL)}{100}} - 0.29$$

Then, $YP = EDGE - MIN$
LIMITS: $0 \leq AZ \leq 360$, $20 \leq SLOPE \leq 45$.

Note that inverse asymmetrical bell-shaped functions, often typical of azimuth effects on plant growth, can be described roughly as above, but in two segments. Each segment then has its own sigmoidal shape. Also, it could easily be that the trough, which was at a fixed point in $AZ$ ($XP = 180$), in the foregoing example, changes in $AZ$ depending on slope. In such a case, $XP = a(SL)$ in the sigmoidal effect of $AZ$.

Concave-Upward Relation in Two Dimensions, with Constant Point of Peaking in X1 Over X2.

Live Data Example

Soil temperatures ($ST$) at various depths and percent of duff reduction were studied on a clearcut area that had been subjected to a controlled burn. A data set (from Shearer 1974 and 1975) was devoted to hypothesis development as shown below.
MATHEMATICAL DESCRIPTION OF GRAPHEd HYPOTHESES - THREE DIMENSIONS

Data were grouped into duff-reduction (DR) classes and trends over "depth in soil" (DS) were smoothed in accord with expectation. These were screened over the reversed DS-scale in the X3-Standards and were all reasonably represented by $bY_7 - DS^p$.

Then, $ST = 45 + (YP + 17) 7 - DS^p$.

And, $YP$, the smoothed curve of scaling heights for the individual DS-curves at the scaling point $[7 - DS = 7]$, was concave upward over $DR$ and could have been described as an exponential function of $DR$. The lower half of a sigmoid, however, was found to mimic the objective curve more closely so,

$$YP = 365.171\left(e^{\frac{-1}{0.56}}\right) - 10.171,$$

thus producing figure 74. Limits of use for this hypothesis are: $0 \leq DR \leq 100$ and $0 \leq DS \leq 7$.

Note that only the lower (left) half of the $YP$-sigmoid is utilized in this relation as the result of doubling the $DR$-scale ($XP = 200$) and the $ST$ scale as well ($YP +$ intercept $= 400$, or $YP = 355$). In the last case, it would have been more consistent to have used $YP = 400$, but a suitable $e^\lambda$ was found with 355; so 355 was retained.

MATHEMATICAL DESCRIPTION OF GRAPHEd HYPOTHESES - FOUR DIMENSIONS

Asymmetrical Bell-Shaped $X1$-Effect with Point of Peaking Moving in $X2$. The $X1$-Effect is Described in Two Segments and Interacts with the Bell-Shaped $X2$-Effect and the Convex-Upward $X3$-Effect (Contrived Example) ........................................... 78

Asymmetrical Bell-Shaped $X1$-Effect, with Constant Point of Peaking in $X2$. The $X1$-Effect is Described in Two Segments and Interacts with Concave-Upward $X2$- and $X3$-Effects (Live Data Example) ........................................... 83

Convex-Upward $X1$-Effect with Peak Moving in $X2$, Interacting with Concave-Upward $X2$-Effects and Slight Convex-Upward $X3$-Effects (Live Data Example) ........................................... 86
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - FOUR DIMENSIONS

Concave-Upward X1-Effects, with Constant Point of Peaking in X2 and Interacting with Convex-Upward X2-Effects; Fourth Dimension Discrete (Live Data Example) .......................................................... 90

Asymmetrical Bell-Shaped X1-Effect with Point of Peaking Moving in X2. The X1-Effect is Described in Two Segments and Interacts with the Bell-Shaped X2-Effect and the Concave-Upward X3-Effect

Contrived Example

Assume that finite curve forms and scales for a 4-D relation have been developed (within gross constraints of the expected relation) from trend information contained in half of a current data set, partitioned over the ranges of the three independent variables. The most prominent feature of the resulting curve array is the asymmetric nature of the X1-effect. This is displayed in figure 75 at three levels of X2, within each of three levels of X3.

Figure 75.—The partial graphed hypothesis.

Also prominent, is the X2-effect, bell shaped, but truncated at the high end of the X2 scale. This is perhaps more fully appreciated in figure 76.

Figure 76.—The complete graphed hypothesis.

The number of X2- and X3-partitions are kept to a minimum to simplify the presentation of associated descriptive methods. More partitions would probably be desirable in a live data application.
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES – FOUR DIMENSIONS

Descriptor methods involved are identical to those shown earlier for 2- and 3-D hypotheses and are simply expanded to accommodate another independent variable. From figure 76 we can anticipate a reasonable mathematical hypothesis to be as shown below.

X1-Effect. — Since the bell-shaped curves are asymmetric, it would be efficient to create a sigmoidal descriptor for the left and right sides separately. They would be joined at X1 and would be scaled to a common height, YP1. Thus, with X1 adjusted to X1 = 50 to accommodate the curve-breadth limit, XP ± XP, the hypothesis for say, the left half, would be:

\[ Y_L = YP1(e^{-k}) \text{ L} \]

or,

\[ Y_L = YP1 \left( \frac{(X1 + 50) - 1}{e^{-1 - IL}} - \frac{1}{e^{-1 - IL}} \right) \]

where YP1, XPI, IL, and nL may change over the range of X2, X3, or both.

YP1. — Plotting the heights of the nine curves from figure 76 in figure 77 and smoothing over these points, YP1 varies as a symmetrical bell-shaped curve (e-~) over X2, changes in height (YP2), and perhaps changes in shape (n and f) over X3. Then, YP1 = YP2(e-~).

Figure 77. — YP1 smoothed.

YP2. — From figure 78, the plotted heights of these curves appear to follow an exponential function of X. Let assume that we have made and compared an overlay curve with the X0-standards and that we have found a suitable descriptor in YP2 = b3(X3)~.

\[ Y_2 = YP2 \left( \frac{(X2 + 10) + 1}{X2} - \frac{1}{X2} - \frac{1}{1 - IL} \right) \]

where

YP2 = b1(X3)~

X2 = 55

n2 = a2 + b2(X3)~

I2 = a2 + b2(X3).

X1 in e-~. — Revision of X1 is necessary, since the smallest XP (from figure 76) approaches zero and the broadest curve reaches X1 = 35 at X3 = 600. Assuming extrapolation to X3 = 700, we might expect the broadest curve to reach no more than say X3 = 50.

To meet the constraint that all curves of the set be within XP ± XP, (X1 = 50) was used in place of X1. Then, where XPI = zero, (X1 = 50) ± 50 and, at the high extreme where X1 = 35, (X1 = 50) = 85. And, using (X1 = 50) in e-~ all curves will be within the minimum XP1 = 50, since ± 2XP1 = 100.

XPI in e-~. — XP1-values read from figure 76 are plotted and smoothed in figure 79.
MATHEMATICAL DESCRIPTION OF GRAPHEd HYPOTHESES – FOUR DIMENSIONS

We will assume by inspection, that $X^P_1$ is an exponential function of $X_2$ and that

$$X^P_1 = (YP_3 - (60)^n(X_2)^n).$$

$YP_3$.—At the scaling point, $X_2 = 60$, the heights ($YP_3$) of the three curves in figure 77 are read as 11, 14, and 18, for $X_3 = 200, 400, 600$, respectively. Plotting these points in figure 80 and smoothing, $YP_3$ can be seen to be a very flat exponential function of $X_3$ so that

$$YP_3 = a + b(X_3)^n.$$  

![Figure 80. YP3 smoothed.](image)

Also, we will assume that the shapes of the three curves in figure 77 have been compared with the $X^n$-Standards and have been found to have a constant $n$.

Then, $X^P_1 = (YP_3 - (60)^n(X_2)^n)$ where $YP_3 = a + b(X_3)^n$.

$nl. and IL, in $e^x$ f $eX$.—We will assume that the left sides of the nine curves in figure 75 have been screened for matching $e^x$ Standards with the results shown in tables 27 and 28.

<table>
<thead>
<tr>
<th>$X_2$</th>
<th>200</th>
<th>400</th>
<th>600</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>40</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>60</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 27. — $nl. in e^x$ $eX$.

<table>
<thead>
<tr>
<th>$X_3$</th>
<th>$X_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td></td>
</tr>
</tbody>
</table>

Table 28. — IL in $e^x$ $eX$.

<table>
<thead>
<tr>
<th>$X_3$</th>
<th>$X_2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>200</td>
<td></td>
</tr>
<tr>
<td>400</td>
<td></td>
</tr>
<tr>
<td>600</td>
<td></td>
</tr>
</tbody>
</table>

Also, we will assume that the left sides of the nine curves in figure 75 have been screened for matching $e^x$ Standards with the results shown in tables 27 and 28.

$IL$.—Referring to table 27, the set of $e^x$-Standards with $n = 2$ was deemed appropriate for the set, all levels of $X_2$ and $X_3$ included. So, $N_2 = 2$, a constant.

$IL$.—The $f$-values in table 28 are quite large since the left-side curves have very steep slopes and are proportionally far to the right in the $(X_1 + 50)$ range. Systematic variation is associated primarily with changes in $X_3$ as indicated by the average $IL$-values. These are plotted and smoothed in figure 81.

<table>
<thead>
<tr>
<th>$X_3$</th>
<th>$IL$</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>0.970</td>
</tr>
<tr>
<td>40</td>
<td>0.965</td>
</tr>
<tr>
<td>60</td>
<td>0.943</td>
</tr>
</tbody>
</table>

Table 28. — IL in $e^x$ $eX$.

![Figure 81. IL smoothed.](image)

And, again assuming this relation can be satisfactorily expressed by an exponential, we have:

$$IL = 0.971 - b(X_3)^n.$$  

Then, to summarize, the entire left side of the hypothesis may be specified as:

$$\hat{Y}_L = YP_1(e^{-x})$$  

or,

$$\hat{Y}_L = YP_1 \left( \frac{(X_1 + 50)}{e^{1 - IL} - 1} - \frac{1}{e^{1 - IL}} \right)$$  

$$\hat{Y}_L = YP_1 \left( \frac{1 - e^{1 - IL}}{1 - e^{-x}} \right)$$

BEST COPY AVAILABLE
As numerical examples, we consider a hypothesized bell-shaped graph with a constant point of peaking in X2.

The hypothesized graph is described in two segments and interacts with concave-upward X3-effects.

In explaining descriptor development here, we focus on generalities. The details of formulating the graphed hypothesis, selecting control points, making overlay curves and comparing them to the standards are assumed to be known from earlier text.

**Mathematical Description of Graphed Hypotheses - Four Dimensions**

Where YPI, XPI, IL, and nL may change over the range of X2, X3, or both.

\[
YP1 = YP2 \left( 1 - \frac{1}{e^\frac{XPI}{1 - I2}} - \frac{XPI}{1 - I2} \right)^{n2} \left( 1 - \frac{1}{e^\frac{IL}{1 - I2}} \right)^{n2}
\]

And

\[
YP2 = b_1 X3^n3
\]

\[
XPI = (10P) \frac{1}{10} A 2^n\text{ and}
\]

\[
YP3 = a + b_1 X3^n3
\]

\[
nL = 2
\]

\[
IL = 0.971 - b_1 X3^n3.
\]

Description of the right side of the hypothesis would be approached in the same fashion as above, to provide the unknowns in:

\[
Y_{RP} = YP1 \left( 1 - \frac{1}{e^\frac{XPI}{1 - I2}} - \frac{XPI}{1 - I2} \right)^{n2} \left( 1 - \frac{1}{e^\frac{IL}{1 - I2}} \right)^{n2}
\]

Note that X1 for the right side curves could be set = (50 - X1) and YPI and XPI are the same for the right side as for the left.

**Asymmetrical Bell-Shaped X1-Effect, with Constant Point of Peaking in X2. The X1-Effect is Described in Two Segments and Interacts with Concave-Upward X2- and X3-Effects**

The relation between total terpene content of lodgepole pine phloem by d.b.h. \(D\) phloem thickness \(P\), and 5-year radial growth \(G\) (from W. E. Cole and others 1981) was hypothesized graphically very nearly as in figure 82.

In explaining descriptor development here, we will focus on generalities. The details of formulating the graphed hypothesis, selecting control points, making overlay curves and comparing them to the standards are assumed to be known from earlier text.

\[
\text{percent terpene content} = \frac{YP1}{(0.14)^{0.35}} P^{0.55}
\]

The heights of the peaks \(YP1\) of the asymmetrical bell-shaped curves over d.b.h. \(D\), at the constant point of peaking \(XPI\), 10.5 inches, are specified as:

For \(D \leq 10.5\)

\[
YP1 = YPD(1.00165e^{\frac{D + 8.5}{19}} - 0.395 - 0.00165)
\]

For \(D > 10.5\)

\[
YP1 = YPD(1.07092e^{\frac{38.5 - D}{28}} - 0.490 - 0.07092)
\]

where

\[YPD = 0.38 + 1.0292(G)^{0.4}\]

Note that sigmoidal curves over \(D\), while having a different shape on opposite sides of the peak, were identified as having consistent shapes for the three growth levels.
MATHEMATICAL DESCRIPTION OF GRAPHAED HYPOTHESES - FOUR DIMENSIONS

This completes the formulation applicable in the range $0 \leq P \leq 0.20$, $0 \leq D \leq 25$, and $0 \leq G \leq 0.35$, but explanation of the $D$-transforms for the sigmoidals may be helpful here. The broadest bell-shaped curve expected at $P = 0.20$ ranges from about $D = -8.5$ on the left side to $D = 38.5$ on the right. With $XP = 10.5$, the $D$ scale for the left side is changed to $(D + 8.5)$ and that for the right is $(38.5 - D)$, as shown in figures 83 and 84.

![Figure 83](image)

**Figure 83.** $D$-transform for the left side.

![Figure 84](image)

**Figure 84.** $D$-transform for the right side.

Again, these transforms of $D$ assure that all points of the sigmoidals are within the range $XP \pm XP$.

The foregoing hypothesis was ultimately refitted to the data from which it was largely derived. This provided the most realistic appraisal of scale available at the moment for the relation characterized.

Given that total terpene estimated from the model = $X$ and that terpene content actually measured in the data = $Y$, the least squares coefficient, $b$, was

$$b = \frac{\Sigma XY - \Sigma X \Sigma Y}{\Sigma X^2} = 1.10301.$$  

This corrects for underscaling in the hypothesis insofar as the data set utilized is concerned. Slight improvement in $R^2$ was noted for the interactive hypothesis over that for the simple additive model with only linear effects of $D$, $P$, and $G$: $R^2 = 0.39$ versus 0.35, respectively. So, no potentially great advantage of the interactive over the additive model appears to exist in this data set. But, the hypothesis still provides new insights for the area of interest that should be of value in planning new studies.

MATHEMATICAL DESCRIPTION OF GRAPHAED HYPOTHESES - FOUR DIMENSIONS

Convex-Upward  
$X_1$-Effect with Peak  
Moving in $X_2$,  
Interacting with  
Concave-Upward  
$X_2$-Effects and Slight  
Convex-Upward  
$X_3$-Effects

Live Data Example

Laboratory-controlled calibration data for Peltier thermocouple psychrometers were assembled by Brown and Bartos, 1981, and are shown smoothed over water potential in figure 85.

![Figure 85](image)

**Figure 85.** The graphed hypothesis and model performance.

The combination of strong expectations for the relation, a well-conceived laboratory study, and strong data-group means left little doubt as to the form and scale of the relation between microvolt output and the independent variables, water potential in negative bars, temperature of the measured medium and the cooling time. The smoothed curves represent an exceedingly strong hypothesis and, with one additional variable, formed the basis for a published mathematical model intended for adjustment to other psychrometers of the same design. The four dimensional relation pictured, in fact, was described mathematically before including the effects of the fifth dimension and provides a unique opportunity to study four dimensional descriptor methods.

The strategy here was to describe microvolt output ($WP$) as a function of water potential ($WP$), temperature ($T$), and cooling time ($CT$), solving the resulting equation for ($WP$).

It was noted at the outset that interest in $WP$ was limited to $XP \leq WP \leq 0$, where $XP$ is the point in $WP$ where microvolts peak. Also, it can be seen in figure 85 that $XP$ changes over both $T$ and $CT$. The heights of the curves, called upper intercepts ($U$) here, change with both $T$ and $CT$, and all curves over $WP$ are oriented at zero. The general form of the relation is estimable from figure 85 as:

$$MV = U - \left(\frac{U}{XP}\right)XP - WP^a.$$  

85
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - FOUR DIMENSIONS

This is an exponential in the reversed WP scale, \( |XP - WP| \), scaled to the difference between \( U/I \) and zero at \( WP = \) zero, and subtracted from \( U/I \). Smoothed values of \( XP \) and \( U/I \) were established and described first, since they were subsequently used in place of actual \( XP \) and \( U/I \)-values to smooth \( XP \)-effects.

\( XP \).—Within each of the three cooling times, the negative \( XP \)-values were read from figure 85 for each of the five temperature curves. The absolute values of these were plotted and smoothed over \( T \) within \( CT \)-groups, with the results shown in figure 86 as

\[
\text{Figure 86} - |XP| \text{ smoothed.}
\]

an exponential effect in the reversed \( T \)-scale, \((40 - T)\), scaled to the difference between an upper intercept \((U/I)\), and scaled to the absolute difference between \( U/I \) and the curve at \( T = 0 \). Then, multiplying by \((-1)\) to return \(|XP|\) to its original negative state,

\[
|XP| = (-1)(U/I - (YP2/(40)^n)(40 - T)^n).
\]

\( U/I \).—The \( U/I \)-values were plotted over \( CT \) and smoothed as shown in figure 87.

\[
\text{Figure 87} - U/I \text{ smoothed.}
\]

This was described as an exponential in the reversed \( CT \)-scale, \((60 - CT)\), subtracted from the upper intercept of 88.0, and scaled to the drop (7.5) at \( CT = 15 \). The resulting form was

\[
U/I = 88 - (7.5/(45)^n)(60 - CT)^n \]
\[
= 88 - 0.0002579 (60 - CT)^n.
\]

\( YP2 \).—Similarly, the \( YP2 \)-values were plotted over \( CT \) and smoothed (fig. 88).

\[
\text{BEST COPY AVAILABLE}
\]

The resulting curve was described as an exponential in the reversed \( CT \)-scale, \((60 - CT)\), scaled to the difference, 1.0, from the intercept at \( CT = 15 \), and added to the intercept. In this case,

\[
YP2 = 8.4 + (1.0/(45)^n)(60 - CT)^n + 0.003475(60 - \text{en}) + 0.0004346(60 - \text{en})^2.
\]

The shapes of the curves over \( T \) in figure 86 were determined through overlay curves and the \( X^n \) standards to be virtually constant at \( n = 2.35 \). Then

\[
XP = (-1)(U/I - YP2/(40^n)(40 - T)^n).
\]

\( U/I \).—Following the analytical pattern for \( XT \), peak values (\( U/I \)) of the 15 curves in figure 85 were read, plotted, and smoothed over \( T \) within each of three \( CT \) groups. Comparison of overlay curves with the \( X^n \) Standards led to adoption of a constant flat form \((n = 1.1)\) for all three curves. Intercepts and scaling heights changed over \( CT \) so that

\[
U/I = INT3 + (YP3/(40)^n)(T)^{1.1}
\]

where

\[
INT3 = 12.1 - 0.0003745(60 - CT)^{1.1}
\]
\[
YP3 = 39.2 - 0.0004936(60 - CT)^{1.1}.
\]

\( n \).—After graphically adjusting the figure 85 curves over WP for smoothed values of \( XP \) and \( U/I \), absolute differences between the \( U/I \) and the associated curves were read for seven points in \( WP \) (\( XP \neq WP \neq 0 \)). The resulting control points were normalized, plotted, and compared to the \( X^n \) Standards. From this, \( n \) varied sigmoidally over both \( T \) and \( CT \) such that

\[
n = 1.18 + 0.185 \left\{ \left[ \frac{40 - T}{40} \right]^n - \left[ \frac{1}{1 - I} \right]^n \right\}.
\]
MATHEMATICAL DESCRIPTION OF GRAPHEO HYPOTHESES - FOUR DIMENSIONS

where

\[ I = 0.45 + 0.000333(T) + (1.9846 \times 10^{-8})(CT)^3 \]

So, with the foregoing inputs and as originally formulated,

\[ MV = \frac{(60 - CT)}{m} \]

where

\[ m = 2.5 + e^{-0.405} \]

The resulting four dimensional model is presented in figure 89 at two points in the fourth dimension (CT = 15, 60).

Data for May and October were subsequently obtained and appeared to be fairly well represented by the July surface. When the hypothesis was fitted to each of the new data sets independently, the resulting surfaces bracketed that for July as expected. The hypothesis was subsequently tested with new, small data sets. Explanations of descriptor development that follow, assume familiarity with 2-D descriptor methods discussed earlier.

The July data were sorted into d.b.h. groups. Within each d.b.h. group, smoothed trends over R were estimated in accord with expectation (shown schematically in figure 91). The negative relation of P to R was strongly displayed in the data, so this was the starting point for the descriptor using the reversed R-scale on an estimated intercept of 0.01. Then

\[ P = 0.01 + (YP(45\degree))50 \cdot R \]
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - FIVE DIMENSIONS

MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - FOUR DIMENSIONS

Figure 91.—The schematic trends.

Screening the resulting curves for n in the $X^2$-Standards, n varied sigmoidally from 1.9 at smaller diameters to about 3.2 at larger ones. The trend was estimated to be as shown in figure 92 where.

$$n = 1.9 + (1.3)e^{0.7}$$

The R-effect was scaled at $R = 5$ or $(50 - R) = 45$, since this was the lower limit of the data over $R$. The height of the July surface at this point minus the intercept of 0.01 = YP.

YP varied positively (convex upward) with D as shown on the July surface (fig. 90) at $R = 5$. This effect (fig. 93) was described as a double component $X^2$-function of the received $D$-scale, subtracted from the uppermost YP. ($YP = 0.158 - 0.010 = 0.148$). The first component was scaled at $D = 4.99$, the second at 0.65. And

$$YP = 0.148 - (2.2687 \times 10^{-3})\{22 - D\}^{22}$$

$$- (1.2694 \times 10^{2}22 - D)^{2}$$

thus completing the inputs for the hypothesis pictured in figure 90. Limits of application are: $0 \leq D \leq 22; 0 \leq R \leq 50$. For $D > 22$, $n = 3.2$ and for $R > 50$, $P = 0.01$.

Based on performance of the hypothesis for the May and October data sets, $R^2 = 0.67$ and 0.78, $r_y = 0.22$ and 0.018, respectively, the July hypothesis appears to represent other months within reason.

Figure 92.—n smoothed.

Figure 93.—YP smoothed.

Discrete-Variable Extension of the 4-D Terpene Example

Page 94

Extension of the 4-D Psychrometer Example (Pages 86-89) to Include the Linear Interactive Effects of $X_3$ and $X_4$
Symmetrical Bell-Shaped X1-Effect with Moving Peak in X3; X2, a Sigmoid, Interacts with X1 to Form the Ridge Shown, Which Changes in Height Sigmoidally Over the Ranges of X3 and X4 (Live Data Example)

Symmetrical Bell-Shaped X1-Effect with Constant Point of Peaking in X2. The Interacting Effect of X2 is Concave-Upward and Those of X3 and X4 are Sigmoidal (Live Data Example)

Extension of the 4-D Psychrometer Example (Pages 86-89) to Include the Linear Interactive Effects of X3 and X4

Live Data Example

The Brown and Bartos (1981) four-dimensional thermocouple example (fig. 89 and 94), ultimately had a fifth dimension. It consisted of the variable, zero-offset microvolts, and is an index of the temperature differential between the thermocouple sensor tips. Such differential has an effect, zero-offset error (ZOE), on the microvolt output that changes depending on the average temperature level of the measured medium and on the direction of the temperature difference between the tips, as indicated in figure 95.
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - FIVE DIMENSIONS

ZOE is quantified in figure 95 only at $CT = 15$ seconds and $WP = -22.5$. A proportional ratio of change, $C$, for microvolts estimated in the four-dimensional model ($MV4$) is derived from $MV$-readings at $WP = -22.5$, as specified below:

$$ C = \frac{MV_z + (MV_z/MV_{15})ZOE}{MV_z} $$

where

$MV_z = MV$ at specified average temperature, $ZOM$, and $15 \leq CT \leq 60$

$MV_{15}$ as above, but with $CT = 15$.

Then,

$MV5 = MV4(C)$

where

$MV4 = f(CT, T, WP)$, as previously specified, and

$ZOE = b_1(ZOM) + b_2(ZOM)(T)$

where

$b_1$ and $b_2$ are estimable directly from figure 96, an annotated version of the right side of the three-dimensional form implied in figure 95.

---

Figure 96.—Zero-offset error in 3-D.

From figure 96,

$$ b_1 = (0.9/60)(ZOM) = 0.015(ZOM) $$

$$ b_2 = (3.4 \times 60 \times 40)(ZOM)(T) = 0.001471(ZOM)(T). $$

Note that the formula for the right side of figure 95 applies to the left side also, since the sign of $ZOM$ and $C$ both become negative there.

This completes the five-dimensional relation, where $MV5 = MV4(C)$ and

$$ MV5 = (U/1 - (U/1)(XP^n)/XP - WP^n)C. $$

Again, with interest focused on $WP$ and solving $MV5$ for $WP$, we have

$$ WP = \left( \left[ \frac{XP^n}{U/1 \times C - MV} \right]^{1/n} + XP \right)^*(-1), \mbox{ for } XP \leq WP \leq 0.
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES - FIVE DIMENSIONS

Discrete-Variable
Extension of the
4-D Terpene Model

Live Data Example

Given the 4-D relation already established for total terpenes (fig. 82 and 97), the residual problem of estimating such a relation for each of three component or component groups of the total (beta-phellandrene, beta-pinene + 3-carene + myrcene, and alpha-pinene) was approached. The associated data were screened graphically over D, P, and G, and were found to contain more relative variability than existed with analysis of total terpenes. There did not appear to be, however, any recognizable departure of basic forms from those estimated to exist for total terpenes.

Then, the obvious alternative of fitting the relation developed for total terpenes to that for each component or component group by least squares was adopted. That is, where estimated total terpene \( T = X \) and measured component terpene \( Y \), the least squares coefficient, \( b \), was

\[
b = \frac{\Sigma XY}{\Sigma X^2}.
\]

The expected sum of the three coefficients independently obtained was the \( b \) of 1.10301 determined for total terpenes (from the 4-D terpene example). So, the three \( b \) values were adjusted by the simple ratio of

\[
\frac{b \text{ for total terpenes}}{10X \text{ for components}}.
\]

Final component coefficients (and associated \( R^2 \)'s) are shown in table 29. Holding \( G \) at its average, 1.59, the resulting surfaces appear as in figure 98.

These surfaces would vary jointly as a rather flat concave-upward curve over \( G \) since, from the 4-D terpene example,

\[
TPD = 0.38 + 1.0292(G)^{-1}.
\]

Figure 97. - The 4-D mathematical hypothesis.

Figure 98. - Discrete partitions for the 4-D terpene hypothesis.

Also, since the form of the relation is identical for each component terpene and the total and since component scales sum to that for the total, components will always sum to the total at any point in the relation.

Symmetrical Bell-Shaped \( X_1 \)-Effect with Moving Peak in \( X_3 \); \( X_2 \), a Sigmoid, Interacts with \( X_1 \) to Form the Ridge Shown, Which Changes in Height Sigmoidally Over the Ranges of \( X_3 \) and \( X_4 \)

Live Data Example

Our hypothesis development here, involves a small data set (from Lee and others 1973) associated with the gap-graded road materials problem discussed under 3-D, asymmetrical bell-shaped ridges. These data were exhausted of information graphically under constraints of expectation to arrive at the mathematical hypothesis shown in figure 99.

Table 29. - 4-D coefficients and \( R^2 \)'s for terpene components

<table>
<thead>
<tr>
<th>Compounds</th>
<th>( b )</th>
<th>( R^2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beta-phellandrene</td>
<td>0.64292</td>
<td>0.402</td>
</tr>
<tr>
<td>Beta-pinene +</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3-carene +</td>
<td>0.3022</td>
<td>0.352</td>
</tr>
<tr>
<td>Myrcene</td>
<td>0.15787</td>
<td>0.213</td>
</tr>
<tr>
<td>Alpha-pinene</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total terpenes (sum)</td>
<td>1.10301</td>
<td>0.387</td>
</tr>
</tbody>
</table>
The explanations of descriptor development that follow presume familiarity with the methods presented in the TWO DIMENSIONS section.

The effect of stability (S) on the dependent variable, ratio (R), was strongly expressed in the data and so was the first effect specified in the descriptor. S appeared to have a consistent sigmoidal shape within partitioned data groupings (some missing because of the small data set) of the remaining three independent variables, flow (F).

\[
\text{RATE} = (YP)c \quad \left( \frac{S}{5000} - 0.9 \right)
\]

This is shown schematically in the top two graphs in figure 100. The starting point in the descriptor is a sigmoid upward to the right over S with zero intercept whose vertical scale, YP, is dependent on the remaining three variables.
The scalar effect of $F$ at $S = 5000$ is bell shaped and symmetric as may be seen in the two graphs (second pair from the top) in figure 100. Adding 20 to $F$ to accommodate the broadest curve of the set under the constraint that $e^{-k}$ functions only between zero and $2XP$, these curves are described by

$$YP = FP \left( \frac{(F+20)}{PXP} - 1 \right) \left( 1 - e^{-\left( \frac{1}{1 - F} \right)^{1.6}} \right)$$

where $FP$, the scalar for $YP$, is estimated (conservatively) to vary sigmoidally with both $V$ and $VF$, as shown in the third tier of graphs from the top of figure 100. The descriptor is

$$FP = FPF + (FPP)e$$

and

$$FPF = 2.4 + 1.8e$$

$$FPP = 0.8 + 2.9e$$

where $FPP$ adds $(2.4 - 2.0 = 0.8)$ at $VF = 100$, and $(7.9 - (2.4 + 1.8 + 0.8) = 2.9)$ at $VF = 0$. $PXP$ is mobile and is assumed to be sigmoidal as shown in the graph at the bottom of figure 100. $PXP$ changes with $VF$ only and can be identified as

$$PXP = 31 + 2.8e$$

And the inflection points ($FI$) for the bell-shaped $YP$-curves vary with $V$, as shown in the left graph of the two just above the one for $PXP$. The graph to the right of that shows associated sigmoidal trends (assumed) over $VF$ for intercepts ($FIF$) and peaks ($FIP$). All told, the descriptor for $FI$ is as follows

$$FI = FIF + (FIP)e$$

where

$$FIF = 0.41 + (0.15)e$$

$$FIP = 0.08 + (0.09)e$$

This chain of descriptor components is converted to the Fortran IV statements shown below:

$$T = \left( \text{EXP}(-\left(\text{ABS}((100 - VF)/40 - 1)/0.42\right)^{0.9}) \right)$$

$$PXP = 31 + 2.8e$$

$$FPF = 2.4 + 1.8e$$

$$FPP = 0.8 + 2.9e$$

$$FIF = 0.41 + 0.15e$$

$$FIP = 0.08 + 0.09e$$

$$PXP = (100 - VF)/40$$

Note that the coefficient of 1.0315 resulted from the refitting of the mathematical model back through the data from which it was derived by least square. Given that modeled rates = $X$ and actual rates = $Y$, $b = \text{EXY}/\text{IX}^2$. Then $bX$ (or $RATE$) is the model submitted to electronic computer/plotter processing that resulted in the graphic presentation of the mathematical hypothesis in figure 99. It is of interest that the additive model with linear defects of $S$, $F$, $V$, and $VF$, reached an $R^2$ of only 0.42 under least squares fit, while, for the model just derived, $R^2 = 0.80$. There would appear to be substantial benefit from utilizing all information at hand for developing the hypothesis in this case.

Symmetrical Bell-Shaped X1-Effect with Constant Point of Peaking in X2.

The Interacting Effect of X2 is Concave-Upward and Those of X3 and X4 are Sigmoidal

Live Data Example

The flexibility of $e^{-k}$ in representing a complex relation is again evident in the five-dimensional interaction pictured in figure 101.
Figure 101. — Deer pellet groups as five-dimensional function of clearcut opening characteristics.
Here, data from Lyon 1976 and Lyon and Jensen 1980 were the basis for an index to intensity of deer use of forest openings created by clearcutting in western Montana. The index is expressed as a function of opening size, height of new vegetation, depth of slash in the opening, and depth of dead and down timber adjacent to the opening.

The data at hand were initially committed to statistical evaluation of the linear effects of the above variables and other independent variables on the intensity index. These were virtually the simplest regression hypotheses that could have been adopted. In this case, the evaluation provided only weak support for expected results and yielded little new information.

To exploit the data further, a hypothesis was developed for the set of four independent variables above. Prior knowledge of the forms of the relations, including interactions, between these variables was summarized. Subject to these constraints, the limited data were then partitioned and graphically exhaust of associated curve form and scaling information. Data appeared to provide strong support for the dynamic interaction anticipated. Procedures specified here were used to develop a functional form, \( XT \), for the graphed interaction. This was refitted to the data by least squares in the model \( PGR = b(XT) \) and \( b = 0.9721 \). The adjusted model is specified below:

\[
PGR = \frac{YP}{(PO - 0.6)^{0.75}} \left( PO - SI \right)^{4.5} (0.9721) \]

where:

- \( PGR \) = number of deer pellet groups per acre inside the opening.
- \( VI \) = height of vegetation in feet, inside the opening.
- \( SI \) = depth of logging slash in feet, inside the opening.
- \( SO \) = depth of dead and down timber, in feet, outside the opening.
- Acres = size of opening in acres.

\[ SI < PO \]

\[ \text{Hypothesis} \]

If \( SI < PO \)

\[
PGR = \frac{YP}{(PO - 0.6)^{0.75}} \left( PO - SI \right)^{4.5} (0.9721) \]

where

\[
YP = YPS \times \left( \frac{\text{Acres} + 200}{1} \right)^{0.75} - e^{-1} \left( \frac{1}{1 - T} \right)^{0.75} + 0.23
\]

\[
YPS = 0.57 + 3.23 \left( \frac{VI}{8 - 1} \right)^{0.76} - e^{-1} \left( \frac{3 - SI}{3 - 1} \right)^{0.4942}
\]

\[
PO = 6.289e^{-1} \left( 0.0999 - 2.289 \right)
\]

Graphic development of the model described by this function apparently resulted in great sensitivity to the interaction information contained in the data. But, since the degrees of freedom thereby sacrificed were unknown, conventional statistical parameters were not estimable. Some indication as to the goodness-of-fit of the functional form to the data set from which it was largely derived is provided by the proportion of the total sum of squares explained by the model, \( R^2 = 0.71 \) here.

Contrast this with the \( R^2 \) of 0.21 achieved with a minimum-effort additive regression model where linear effects of the same four independent variables have been fitted to the data by least squares. It would appear that sharp focus on interaction formulation was justified.

At this point in the paper, it is assumed that only a minimum of explanatory information will be necessary to clarify descriptor development.

The expected negative SI effect was strongly apparent in the data and so was used as the starting point for the descriptor. This effect was exponential over the reversed SI-axis where \( 0.6 < SI < PO \). The lower extreme of the data in SI, 0.6, was adopted as the scaling point (largest PGR-values) in the reversed SI-scale. PO, a variable over VI, represented the point in SI where the PGR-curves reached zero (fig. 102).

When the point of exponential origin was allowed to vary, \( n \) in \( X^n \) approached a constant (1.55) over SI, but the scaling heights of these curves (\( 0 < SI < PO \) varied over ACRES, VI and SO. Then,

\[
PGR = (50 \times b) \left( \frac{YP}{(PO - 0.6)^{0.75}} \right) \left( PO - SI \right)^{4.5} , SI < PO
\]
MATHEMATICAL DESCRIPTION OF GRAPHED HYPOTHESES – FIVE DIMENSIONS

where

50 is the constant necessary to the conversion of data units to acres;

b is the least squares coefficient fitting the derived model back through the source data.

Note that segmentation is employed in the exponential variable here, since beyond $PO$ in $Acres$, $PO$ is specified as zero. $PO$ varied sigmoidally over $1/V$ and

$$PO = \frac{1}{0.999} \cdot \left(1 + 10^{\frac{1}{0.5}} \cdot e^{-10^{\frac{8}{1.75}}} \right)$$

$YP = 0.23 + 3.23 \cdot e^{-0.404 \cdot 1.75}$. $XP$ appeared constant at 260 acres and, from the $e^{-A}$-Standards, $n = 1.75$ was reasonably representative for all curves. $I$, however, varied sigmoidally over $1/V$ and

$$I = 0.897 - 0.067 \cdot e^{-0.404 \cdot 3.5}$$

$YPS$, the heights of the bell-shaped curves over $ACRES$, varied as an interaction of sigmoids, a new descriptor feature. $YPS$ varied sigmoidally over $1/V$ as

$$YPS = 0.57 + 3.23 \cdot e^{-0.76 \cdot 1.75}$$

but only achieved this magnitude of effect at smaller $SO$-values. The $1/V$-effect became sigmoidally smaller as $SO$ increased. Noting that the unscaled sigmoid over the reversed $SO$-axis ($3 - SO$) ranges in value from zero to one, the $SO$-effect,

$$\left(3 - \frac{SO}{3}\right)^{10} \cdot e^{-0.492}$$

becomes a convenient multiplier with which to achieve the appropriate sigmoidal reduction of the maximum $1/V$-effect.

All curve form estimates were derived from data-based graphic hypotheses, normalized control points, overlay curves, and comparison to either the $X^e$- or $e^{-A}$-Standards. While

the foregoing mathematical specifications may seem voluminous, they are quite easily specified in a short Fortran IV program for computer processing, as shown:

```fortran
PO = 6.289 \cdot (EXP(-ABS((1/V10 - 1)/0.999)) - 2.289)
A = EXP(-ABS((I - SO))/(3 - 1/0.492))
B = EXP(-ABS((1/V8 - 1)/0.76))
YPS = 3.23 \cdot B + 57
I = 0.897 - 0.067 \cdot EXP(-ABS((1/V8 - 1)/0.404))
YPS = 0.57 + 3.23 \cdot e^{-0.76 \cdot 1.75}
YP = 0.23 + 3.23 \cdot e^{-0.404 \cdot 1.75}
PGR = (YP \cdot YPS \cdot (LN - RN)/(1 - RN) + 0.23)
X(5) = PGR
RETURN
1 X(5) = 0
RETURN
END
```

Note that figure 102 was plotted electronically using a program that incorporated the above statements.
PUBLICATIONS CITED

Bartlett, M.S.


Damaerschalk, J. P., and A. Kozak.


Draper, N. R., and H. Smith.

George, C. W., and A. D. Blakely.

Jensen, C. E.

Jensen, C. E.

Jensen, C. E.

Jensen, C. E., and J. W. Homeyer.

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

Kars, 0. J.

ADDITIONAL REFERENCES

Lee, D-Y., H. T. David, and R. W. Mensing.

Lyon, L. J.

Lyon, L. J., and C. E. Jensen.

Mead, R., and D. J. Pike.

Mosteller, F., and J. W. Tukey.

ADDITIONAL REFERENCES


Basile, J. V., and C. E. Jensen.

Clary, W. P., and C. E. Jensen.

Clayton, J. L., and C. E. Jensen.

Cleendener, G. W.

Cole, W. E.


George, C. W.
APPENDIX A

APPENDIX-A: $e^A$ DERIVATION

The new function, simply identified as $e^A$, provides the analyst with a finite source of versatile transformations for use when it is inefficient to search for alternatives or when alternatives are inadequate. While $e^A$ is not a panacea for all the problems models face in developing functional relationships, it along with curves of the class $X^n(\pi \pm 1)$ serves a broad spectrum of transformation needs with no particular limitations as to curve shapes for which it is most useful. Methods for developing mathematical descriptors using these functions have been treated by Jensen and Homeyer (1970, 1971) and Jensen and Jenson (1973, 1976, and 1979) and are summarized and updated in this paper.

A variant of the Normal function $e^A$ is defined as:

$$e^A = \left(1 - \frac{(X/XP) - 1}{(1 - l)}\right)^a - \left(1 - \frac{(X/XP) - 1}{(1 - l)}\right)^b, \ 0 \leq X \leq 2XP$$

where,

- $e$ = natural logarithm base.
- $XP$ = pivot point in $X$ for $e^A$ and is the point in $X$ at which the bell-shaped curve peaks.
- $XI$ = point of sigmoidal inflection in $X$.
- $l$ = proportional departure of $XI$ from zero to $XP$ or from $2XP$ to $XP$.
- $n$ = power of negative exponents for $e$.

The left numerator, like the Normal function, generates a system of bell-shaped curves about a central point in $X, XP$ in this case. These curves reach the maximum value of 1 at $XP$ and then decline sigmoidally and symmetrically on either side of $XP$ with increasing departure of $X$ from $XP$.

User control of the $e^A$ system is enhanced if all curves of the set range in value from zero to one within a finite domain of $X$. To achieve this property, curves generated by the left numerator have been forced through zero at $XP + XP$ as follows: Let the left numerator be $e^{-l}$ and let $e^{-l} = e^{-l}$ at $X = 0$ or $2XP$. Consider the residuals generated by $1 - e^{-l}$. Expanded by the inverse $(1 - e^{-l})^{-1}$ and subtracted from one, we have

$$e^A = 1 - (1 - e^{-l})^{-1}(1 - e^{-l})$$

which simplifies to the final form

$$e^A = (e^{-l})^{-1} = e^{-l}$$

Then the right numerator and the denominator serve to force curves of the left numerator through zero at $XP + XP$. This domain, along with $XP$, can be altered to accommodate the needs of individual conceptual models by adding constants $(+ \ or -)$ to the $X$-scale. The apparent complexity of $e^A$ is much reduced by the fact that in application, the right numerator and the denominator simplify to constants or approach zero and can be deleted.
APPENDIX A (con.)

Sigmoid s on either side of $X_P$ are forced into different areas of two-dimensional space by shifting $I$, the proportional point of inflection in $X$, and the slope of the sigmoidal face at $I$ changes with $n$.

Arrays of curves from the left half of this function (a mirror image of the right half) are shown in the $e^{-x}$-Standards that follow. Sets are given for commonly used increments of $n$, $1.2 \leq n \leq 10.0$. Each set has curves progressing from $I = 0.1$ at the left to 0.9 at the right. These curves or any portions thereof provide an almost endless potential for matching graphed curves and describing them mathematically. Given that $Y_P$ is the peak value of $Y$ (the dependent variable) or the objective curve, the selected $e^{-X}$-function ($e^x$ ranges from zero to one) may be scaled to that curve through multiplication by $Y_P$.

While the controllable parameters $I$, $n$, $X_P$, and $Y_P$ provide exceptional descriptive power in two dimensions, such power is magnified manyfold for hypotheses involving three or more dimensions, as well as strong interactions. For example, the location of peaks $Y$ that move across the independent variables in any fashion can be described with $X_P = f(I(X))$. The heights of these peaks in any pattern and changing sigmoidal shapes leading up to them can again be pictured as functions of the independent variables $Y_P = f_2(X)$, $n = f_3(X)$, and $I = f_4(X)$. Additionally, asymmetry in bell-shaped hypotheses is easily accounted for by describing sigmoids on either side of $X_P$ independently (segmented descriptors), while scaling both to the same height at $X_P$.

A descriptor $(X_T)$, adopted as a hypothesis in its entirety, may be fitted to pertinent data by least squares in the simple model: $Y = \beta X_T + \epsilon$, where $\epsilon$ are $NID(0, \sigma^2)$, constant variance, and $\beta = \Sigma X_T Y / \Sigma X_T^2$.

Weighted regression procedures are recommended where the variance about regression is not uniform over the ranges of the independent variables. In such cases, reasonable success in achieving constant variance has been obtained by solving for departures $Y - Y$ or $d^2$ as a suitable function of the related $Y$ in $d^2 = b Y$. Then the weights, $w$, for each observation is set equal to $1/\sigma^2$ and a weighted $\beta$ in $Y = \beta Y$ is estimated as $\beta = \Sigma w Y / \Sigma w$.

More complex, generally iterative fitting procedures, such as the Newton–Raphson methods, can be used to arrive at statistical estimates of internal model parameters (see Damaerschalk and Kozak 1977; Draper and Smith 1966).
$N = 1.2$
\[ N = 1.4 \]
\[ N = 1.6 \]

Graph showing the relationship between \( Y/YP \) and \( X/XF \) for various values of \( I \).
$N = 1.8$
$N = 2.0$

Graph showing $Y/YP$ vs $X/XP$ for various values of $I$. The graph is for $N = 2.0$. The axes range from 0 to 1.
N = 2.2

\[ \frac{Y}{Y_P} \]

\[ \frac{X}{X_P} \]
\[ N = 2.6 \]
$N = 2.8$
\[ N = 3.2 \]
\[ N = 3.4 \]
$N = 3.6$

Diagram showing curves for $Y/YP$ and $X/XP$ with $N = 3.6$.
\[ Y/YP \]

\[ X/XP \]

\[ N = 4.0 \]
N = 6.0
\[ N = 7.0 \]
\[ N = 8.0 \]
The problem of over-simplified regression hypotheses is identified and alternatives are discussed. Methods for developing structured hypotheses, first graphically and then mathematically, are shown for two through five-dimensional cases. Mathematical formulation is accomplished through use of two families of curves, $X^n, n \geq 1$, and $e^{-x}$, a highly manipulable function for bell-shaped curves or portions thereof.

KEYWORDS: structured regression hypotheses, regression hypotheses, modeling, interactive models, multidimensional descriptive geometry