By Stanford L. Arner

## -ROFIL

A computer prosrom for disployng the geonetric relothonshlps of the responser in o foctorial desion

## USDA FOREST SERVICE RESEARCH PAPER NE-288 <br> 1974

NORTHEASTERN FOREST EXPERIMENT STATION
FOREST SERVICE U.S. DEPARTMENT OF AGRICULTURE
6816 MARKET STREET, UPPER DARBY, PA. 19082
WARREN T. DOOLITTE DIRECTOR

## The Author

STANFORD L. ARNER has served as a computer programmer and mathematical statistician in the biometrics branch of the Northeastern Forest Experiment Station since 1966. He received a Bachelor of Science degree in Forest Science from the Pennsylvania State University in 1966 and a Master of Science degree from the School of Forest Resources at the Pennsylvania State University in 1971.

## PROFILE

## A computer program for displaying the geometric relationships of the responses in a factorial design


#### Abstract

One procedure in analyzing the responses to factors in an experiment is to construct profiles of the means. Profiles are plots of the sample means of each level of one factor at each level of one or more of the other factors in the experiment. The construction of these profiles and tests of hypotheses about the mean responses is called profile analysis. This paper discusses profile analysis and presents a computer program which produces profiles of the experimental means.


ONE OBJECTIVE of designed experiments is to find the relationship of the response to various factors in the experiment. Profile analysis is a technique for studying the geometric relationships among the expected responses in a multifactor experiment. Profiles are line segments connecting the means for the levels of one factor (A) at each level of a second factor (B). If there is a third factor (C), there are three possible simple profiles: the AB profiles at each level of C; the $A C$ profiles at each level of $B$; and the BC profiles at each level of A. Also, there are the two factor profiles combined over all levels of the third factor, and the one factor profiles combined over all levels of the other two factors.
To introduce profile analysis, an example of a three-way factorial experiment is presented. In this hypothetical experiment the height growth of two species of trees fertilized with two levels of nitrogen and three levels of phosphorus is measured. Let $Y_{i j k m}$ be the height growth on the $\mathrm{m}^{\text {th }}$ tree of the $\mathrm{i}^{\text {th }}$ species (factor A), fertilized with the $\mathrm{j}^{\text {th }}$ level of nitrogen (factor B) and the $\mathrm{k}^{\text {th }}$ level of phosphorus
(factor C). The mean height growth of the $\mathrm{M}_{\mathrm{ijk}}$ trees in the $\mathrm{ijk} \mathrm{k}^{\mathrm{th}}$ factor combination is

$$
\bar{Y}_{i j k}=\sum_{m=1}^{M_{i j k}} Y_{i j k m} / M_{i j k} .
$$

Experimental values for the $\overline{\mathrm{Y}}_{i j \mathrm{k}}$ are presented in table 1. For demonstration purposes

Table 1.-Means for simple profiles

|  | $\mathbf{A}_{1}$ |  |  |  |  | $\mathbf{A}_{2}$ |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathbf{C}_{1}$ | $\mathbf{C}_{2}$ | $\mathbf{C}_{2}$ |  | $\mathbf{C}_{1}$ | $\mathbf{C}_{2}$ | $\mathbf{C}_{3}$ |  |
|  | $\mathbf{B}_{1}$ | 40 | 30 | 60 |  | 20 | 10 |  |
| $\mathbf{B}_{2}$ | 60 | 50 | 20 |  | 40 | 30 | 0 |  |

we assume that these are population values. With real data the values obtained would be sample means rather than population means. The expected value of the sample means would be the population means; that is,

$$
\mathrm{E}\left(\overline{\mathrm{Y}}_{\mathrm{ijk}}\right)=\mu_{\mathrm{ijk}} .
$$

The simple profiles of interest might be the nitrogen by phosphorus profiles for each species. One of the nutrient factors (say phosphorus) is chosen for the horizontal axis, the
other (nitrogen) is chosen for the vertical axis. The 3 levels of phosphorus are then the points along the horizontal axis at which the mean responses $\overline{\mathrm{Y}}_{i j k}$ are plotted for each level of nitrogen. Using the values in table 1, $\overline{\mathrm{Y}}_{111}$ would be positioned at a height of 40 on the vertical axis directly above the first index point of the horizontal axis; $\overline{\mathrm{Y}}_{121}$ would be directly above the first point at 60 on the vertical axis. The simple profiles for species 1 are the first set of profiles in figure 1, while those for species 2 are the second set.
The means for the two-way profiles combined over the third factor are:

$$
\begin{aligned}
& \overline{\mathrm{Y}}_{. j \mathrm{k}}=\sum_{\mathrm{i}=1}^{2} \overline{\mathrm{Y}}_{\mathrm{ijk}} / 2 \\
& \overline{\mathrm{Y}}_{\mathrm{i} \cdot \mathrm{k}}=\sum_{\mathrm{j}=1}^{2} \overline{\mathrm{Y}}_{\mathrm{ijk}} / 2 \\
& \overline{\mathrm{Y}}_{\mathrm{ij} .}=\sum_{\mathrm{k}=1}^{3} \overline{\mathrm{Y}}_{\mathrm{ijk}} / 3
\end{aligned}
$$

For example:

$$
\overline{\mathrm{Y}}_{.11}=\left(\overline{\mathrm{Y}}_{111}+\overline{\mathrm{Y}}_{211}\right) / 2=(40+20) / 2=30
$$

Figure 1.-Simple BC profiles for each level of factor $A$.



Table 2.-Means for two-way profiles combined over third factor

| $\mathrm{A}_{1}+\mathrm{A}_{2}$ |  |  |  | $\mathrm{B}_{1}+\mathrm{B}_{2}$ |  |  |  | $\mathrm{C}_{4}+\mathrm{C}_{2}+\mathrm{C}_{3}$ |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  | $\mathrm{C}_{1}$ | $\mathrm{C}_{2}$ | $\mathrm{C}_{3}$ |  | $\mathrm{C}_{1}$ | $\mathrm{C}_{2}$ | $\mathrm{C}_{3}$ |  | $\mathrm{A}_{1}$ | $\mathrm{A}_{2}$ |
| $\mathrm{B}_{1}$ | 30 | 20 | 50 | $\mathrm{A}_{1}$ | 50 | 40 | 40 | $\mathrm{B}_{1}$ | 43.33 | 23.33 |
| $\mathrm{B}_{2}$ | 50 | 40 | 10 | $\mathrm{A}_{2}$ | 30 | 20 | 20 | $\mathrm{B}_{2}$ | 43.33 | 23.33 |

Figure 2.-Two-way profiles combined over third factor.


The values of the means combined over the third factor are given in table 2 , and the combined profiles are given in figure 2.

The last sets of profiles are based on the means combined over two factors. These are:

$$
\begin{aligned}
& \bar{Y}_{. . k}=\underset{j=1}{2} \bar{Y}_{. j k} / 2=\sum_{i=1}^{2} \bar{Y}_{i . k} / 2 \\
& j=1 \quad i=1 \\
& \overline{\mathrm{Y}}_{\mathrm{j} .}=\underset{\mathrm{i}=1}{2} \overline{\mathrm{Y}}_{\mathrm{ij} .} / 2 \underset{\mathrm{k}=1}{\sum_{\mathrm{Y}}} \overline{\mathrm{Y}}_{\mathrm{jk}} / 3
\end{aligned}
$$

$$
\overline{\mathrm{Y}}_{\mathrm{i} . .}=\sum_{\mathrm{j}=1}^{2} \overline{\mathrm{Y}}_{\mathrm{ij} .} / 2=\sum_{\mathrm{k}=1}^{3} \overline{\mathrm{Y}}_{\mathrm{i} . \mathrm{k}} / 3
$$

The values of the means combined over two factors are given in table 3 and the combined profiles are shown in figure 3.

Table 3.-Means for single factors combined over other two factors

| $\mathbf{B}_{1}$ | 33.33 | $\mathrm{~A}_{1}$ | 43.33 | $\mathrm{C}_{1}$ | 40 |
| :--- | :--- | :--- | :--- | :--- | :--- |
| $\mathbf{B}_{2}$ | 33.33 | $\mathrm{~A}_{2}$ | 23.33 | $\mathrm{C}_{2}$ | 30 |
|  |  |  | $\mathbf{C}_{3}$ | 30 |  |

Figure 3.-Profiles on one factor combined over other two factors.


## Profile Analysis <br> And The Factorial Model

The statistical model for a three way factorial design can be written

$$
\begin{aligned}
\mu_{\mathrm{ijk}}=\mu & +\mathrm{A}_{\mathrm{i}}+\mathrm{B}_{\mathrm{j}}+\mathrm{C}_{\mathrm{k}}+\mathrm{AB}_{\mathrm{ij}}+\mathrm{AC}_{\mathrm{ik}} \\
& +\mathrm{BC}_{\mathrm{jk}}+\mathrm{ABC}_{\mathrm{ijk}}
\end{aligned}
$$

further

$$
\begin{aligned}
& \Sigma_{\mathrm{i}} \mathrm{~A}_{\mathrm{i}}=\Sigma_{\mathrm{j}} \mathrm{~B}_{\mathrm{j}}=\Sigma_{\mathrm{k}} C_{\mathrm{k}}=0 \\
& \Sigma_{\mathrm{i}} \mathrm{AB}_{\mathrm{ij}}=\Sigma_{\mathrm{j}} \mathrm{AB}_{\mathrm{ij}}=\Sigma_{\mathrm{i}} \mathrm{AC}_{\mathrm{ik}}=\Sigma_{\mathrm{k}} \mathrm{AC}_{\mathrm{ik}}=\Sigma_{\mathrm{j}} \mathrm{BC}_{\mathrm{jk}} \\
&=\Sigma_{\mathrm{k}} \mathrm{BC}_{\mathrm{jk}}=0 \\
& \Sigma_{\mathrm{i}} \mathrm{ABC}_{\mathrm{ijk}}=\Sigma_{\mathrm{j}} \mathrm{ABC}_{\mathrm{ijk}}=\Sigma_{\mathrm{k}} \mathrm{ABC}_{\mathrm{ijk}}=0
\end{aligned}
$$

The parameters are the main effects $\mathrm{A}, \mathrm{B}, \mathrm{C}$; the two-way interactions $\mathrm{AB}, \mathrm{AC}, \mathrm{BC}$; and the three-way interaction $A B C$.

For interpretation of the two-way interactions we look at the two-way profiles combined over the levels of the third factor. For example, the means combined over factor A are

$$
\begin{gathered}
\mu_{. j \mathrm{k}}=\sum_{\mathrm{i}=1}^{\mathrm{I}} \mu_{\mathrm{ijk}} / \mathrm{I} \\
=\left(\mathrm{I} \mu+\Sigma \mathrm{A}_{\mathrm{i}}+\mathrm{I} \mathrm{~B}_{\mathrm{j}}+\mathrm{IC} \mathrm{C}_{\mathrm{k}}+\Sigma \mathrm{AB}_{\mathrm{ij}}+\right. \\
\left.\Sigma \mathrm{AC}_{\mathrm{ik}}+\mathrm{IBC}_{\mathrm{jk}}+\Sigma \mathrm{ABC}_{\mathrm{ijk}}\right) / \mathrm{I} \\
=\mu+\mathrm{B}_{\mathrm{j}}+\mathrm{C}_{\mathrm{k}}+\mathrm{BC}_{\mathrm{jk}},
\end{gathered}
$$

since the restrictions imposed on the parameters result in all terms involving A summing to zero.

The $\mu_{\mathrm{jk}}$ 's are the means plotted on the combined BC profiles which are used to determine if there is a BC interaction. The interaction is the degree to which the profiles are not parallel. If the profiles are parallel there is no interaction.

The BC interaction can be expressed in terms of $\mu_{\mathrm{j}, \mathrm{k}}$. If there is no BC interaction the distance between profiles is the same at each point along the horizontal axis. Therefore $\mu_{1 \mathrm{k}}-\mu_{2 \mathrm{k}}$ is constant for all k. For this example the hypothesis of no BC interaction is:

$$
\begin{aligned}
& \left(\mu_{.11}-\mu_{.21}\right)-\left(\mu_{.12}-\mu_{.22}\right)=0 \\
& \left(\mu_{.12}-\mu_{.22}\right)-\left(\mu_{.13}-\mu_{.23}\right)=0
\end{aligned}
$$

These linear combinations of means are functions of the BC interactions only since

$$
\begin{aligned}
& \mu_{.11}-\mu_{.21}-\mu_{.12}+\mu .22 \\
& =\left(\mu+\mathrm{B}_{1}+\mathrm{C}_{1}+\mathrm{BC}_{11}\right)-\left(\mu+\mathrm{B}_{2}+\mathrm{C}_{1}+\mathrm{BC}_{21}\right) \\
& -\left(\mu+\mathrm{B}_{1}+\mathrm{C}_{2}+\mathrm{BC}_{12}\right)+\left(\mu+\mathrm{B}_{2}+\mathrm{C}_{2}+\mathrm{BC}_{22}\right) \\
& =\mathrm{BC}_{11}-\mathrm{BC}_{21}-\mathrm{BC}_{12}+\mathrm{BC}_{22} .
\end{aligned}
$$

Likewise,

$$
\begin{aligned}
& \mu_{.12}-\mu_{.22}-\mu_{.13}+\mu_{.23}= \\
& \mathrm{BC}_{12}-\mathrm{BC}_{22}-\mathrm{BC}_{13}+\mathrm{BC}_{23} .
\end{aligned}
$$

For this example (table 2),

$$
\begin{aligned}
& \left(\mu_{.11}-\mu_{.21}\right)-\left(\mu_{.12}-\mu_{.22}\right)= \\
& (30-50)-(20-40)=0 \\
& \left(\mu_{.12}-\mu_{.22}\right)-\left(\mu_{.13}-\mu_{.23}\right)= \\
& (20-40)-(50-10)=-60
\end{aligned}
$$

implying that the BC interaction is not zero. The BC profiles (fig. 2), show that the profile segments between $\mathrm{C}_{1}$ and $\mathrm{C}_{2}$ are parallel and the segments $\mathrm{C}_{2}$ and $\mathrm{C}_{3}$ are not parallel.

The combined AB and AC profiles provide similar information about the AB and AC interactions. In the example, both the $A B$ and AC profiles are parallel. Therefore there is no AB or AC interaction. The interactions associated with the combined profiles are called the overall interaction effects.

Now consider the simple BC profiles for one level of A, (fig. 1). The profiles are not parallel; that is, the BC interaction at $a_{1}$ and $a_{2}$ is not zero. These interactions are called simple interaction effects (denoted BC for $a_{i}$ ).

When each cell of the experiment has the same number of observations, the variation in the simple interactions is related to the variation in the overall interactions as follows:

$$
\mathrm{SS}_{\mathrm{BC} \text { at } \mathrm{a}_{1}}+\mathrm{SS}_{\mathrm{BC} \text { at } \mathrm{a}_{2}}=\mathrm{SS}_{\mathrm{BC}}+\mathrm{SS}_{\mathrm{ABC}} .
$$

In the example the variations due to the simple BC interactions are both positive, therefore one or both of $\mathrm{SS}_{\mathrm{BC}}$ and $\mathrm{SS}_{\mathrm{ABC}}$ are positive. We have seen that $\mathrm{SS}_{\mathrm{BC}}$ is not zero. But what is the value of $\mathrm{SS}_{\mathrm{Abc}}$ ?

For information about the variation due to the overall ABC interactions, we compare the
simple BC profiles. The three-way interaction is zero when the BC interactions for each level of A are similar. This could occur either when the simple profiles are parallel to each other or when both simple interactions are zero. For our data, the individual simple profiles are not parallel, but the sets are parallel to each other. Therefore the ABC interaction is zero.

In terms of means, the hypothesis of no three-way interaction is,

$$
\left[\begin{array}{l}
\left(\mu_{111}-\mu_{121}\right)-\left(\mu_{112}-\mu_{122}\right)= \\
\left(\mu_{211}-\mu_{221}\right)-\left(\mu_{212}-\mu_{222}\right) \\
\left(\mu_{112}-\mu_{122}\right)-\left(\mu_{113}-\mu_{123}\right)= \\
\left(\mu_{212}-\mu_{222}\right)-\left(\mu_{213}-\mu_{223}\right)
\end{array}\right]
$$

or,

$$
\left[\begin{array}{l}
\left(\left(\mu_{111}-\mu_{121}\right)-\left(\mu_{112}-\mu_{122}\right)\right)- \\
\left(\left(\mu_{211}-\mu_{221}\right)-\left(\mu_{212}-\mu_{222}\right)\right)=0 \\
\left(\left(\mu_{112}-\mu_{122}\right)-\left(\mu_{113}-\mu_{123}\right)\right)- \\
\left(\left(\mu_{212}-\mu_{222}\right)-\left(\mu_{213}-\mu_{223}\right)\right)=0 .
\end{array}\right]
$$

For our example the values for the hypothesis are,

$$
\left[\begin{array}{l}
((40-60)-(30-50))- \\
((20-40)-(10-30))=0 \\
((30-50)-(60-20))- \\
((10-30)-(40-0))=0
\end{array}\right]
$$

The hypothesis of no three-way interaction is a function of the ABC interaction only and can be expressed in these terms by substituting $\mathrm{ABC}_{i j \mathrm{k}}$ for $\mu_{\mathrm{ijk}}$.

The interpretation of mean profiles can be extended to experiments with more than three factors. For instance, the four-way profiles could be created by plotting the two factor profiles at each combination of levels of the other two factors. Interpretation of parameters and tests of the statistical model for experiments involving more than three factors is difficult, however.

Table 4 summarizes the relationships that the sums of squares for the two-way and three-way interactions have with the form of the simple and combined two-way profiles. As an example, suppose the test results given

Table 4.-Relationship of simple and combined profiles to the variation due to twoand three-way interactions

| Lines in combined proflle | Lines in one simple profile compared to lines in all other simple profiles | Lines in each simple profile | Value of |  |
| :---: | :---: | :---: | :---: | :---: |
|  |  |  | $\mathrm{SS}_{4 B}$ | SS ${ }_{\text {ABO }}$ |
| NP | NP | NP | $>0$ | $>0$ |
| NP | P | NP | $>0$ | 0 |
| P | NP | NP | 0 | $>0$ |
| P | P | P | 0 | $\bigcirc$ |
| P | NP | P | 0 | 0 |

in the last line of table 4 are encountered. If the lines of the $A B$ profile combined over factor C are parallel and the lines in each simple $A B$ profile are parallel, even though the simple profiles are not parallel to each other, both $\mathrm{SS}_{\mathrm{AB}}$ and $\mathrm{SS}_{\mathrm{ABC}}$ are zero.

For a more detailed discussion of profile analysis, see Winer 1962, pp. 162-184.

## DESCRIPTION OF THE COMPUTER PROGRAM-PROFILE

Program Profile is designed to show the relationships between the response and various factors in the experiment.

Any pair of factors can be chosen for the vertical or the horizontal axis. Further, a third factor can be designated such that profiles will be created for each level of this third factor. Also, specific levels of any of the factors can be selected so that a set of profiles is plotted for the one level only. If there are multiple observations for each level of the factors - for example, measurements on a variable at different time periods-the response vector can be designed as a factor using the different observations as the levels.
The scale for the levels of the factor on the horizontal axis can have unequal distances between levels. If the time periods are 1,2 , and 4, for example, the distance between levels 2 and 3 is twice the distance between levels 1 and 2.

The responses can be transformed by in-
sertion of a user-written FORTRAN subroutine.

The program is designed to handle data from multiple factor experiments. If there is a subclass for which no record is observed, then that cell is considered missing from the design. A missing cell is ignored in the calculation of a combined mean. Therefore the profile of the combined means must be interpreted with caution.

There is an option that is useful when the response variables are counts. In this case the absence of records implies a zero count for that cell. The cell is not considered missing from the design, and is therefore included in the calculation of a combined mean. If the design is incomplete, the missing cells can be identified by the user.

The differences between the two methods can best be illustrated by the calculation of cell means. Suppose there are two factors, each with two levels. If there is at least one observation in each cell the means are calculated as follows:

Normal Procedure

## Counts Option

$\bar{Z}_{i j}=Z_{i j} / n_{i j}$

$$
\overline{\mathrm{Z}}_{\mathrm{ij}}=\mathrm{Z}_{\mathrm{ij}} / 1
$$

$$
\overline{\mathrm{Z}}_{\mathrm{i} .}=\left(\overline{\mathrm{Z}}_{\mathrm{i} 1}+\overline{\mathrm{Z}}_{\mathrm{i} 2}\right) / 2 \quad \overline{\mathrm{Z}}_{\mathrm{i} .}=\left(\mathrm{Z}_{\mathrm{i} 1}+\mathrm{Z}_{\mathrm{i} 2}\right) / 2
$$

If there is no record for $\mathrm{Z}_{12}$ and $\mathrm{Z}_{12}$ is not designed as missing for the counts option;

$$
\begin{array}{ll}
\overline{\mathrm{Z}}_{1 .}=\overline{\mathrm{Z}}_{11} / 1 & \overline{\mathrm{Z}}_{1 .}=\left(\mathrm{Z}_{11}+0\right) / 2 \\
\overline{\mathrm{Z}}_{2}=\left(\overline{\mathrm{Z}}_{21}+\overline{\mathrm{Z}}_{22}\right) / 2 & \mathrm{Z}_{2}=\left(\mathrm{Z}_{21}+\mathrm{Z}_{22}\right) / 2
\end{array}
$$

## Program Control Cards

## 1. Data Set Control Card Column

1, 2 NS $=$ Number of factors in design.
3,4 NYR $=$ Number of responses per observation (length of response vector).
$5,6 \mathrm{NF}=$ Number of format cards $(\leq 3)$.
$7,8 \mathrm{NFL}=$ Input file number ( 5 is default).
9, 10 PM If $\mathrm{PM}=0$ print all cell means of complete design.
If $\mathrm{PM}=1$, skip.

11, 12 NT If NT $=0$ no data transformation.
If $\mathrm{NT}=1$ transform data with user-written subroutine.
13, 14 ICNT If ICNT $=0$ responses are measurements on some continuous variable.
If $\operatorname{ICNT}=1$ responses are frequency counts.
15,16 NDES $=$ Number of cells not included in design. (If ICNT=1 and some of the factor combinations are not included in the design of the experiment NDES is the number of cards to be read containing the factor levels for which the observation count is set to zero.)
21-25 LEVELS (1) = Number of levels of response factor after transformation.
26-75 LEVELS $(\mathrm{L})=$ Number of levels of factor L ; $\mathrm{L}=1$ to NS). These are consecutive fields of length 5.

## 2. Variable Format Cards

These begin with a left parenthesis in column 1 of the first card which is followed by NS I format designations and NYR F or E format designations. A right parenthesis should follow the last F or E format field.

## 3. Missing Cell Cards (Optional)

One card for each missing criterion and/or cell in the design. The first 2 digit code designates the element of the response vector and the remaining 2 digit codes represent the factor levels of the missing values; card format (11 I2). Missing cell cards are needed when $\mathrm{ICNT}=1$ and $\mathrm{NDES} \neq 0$.

## 4. Data Cards

The data card format must conform to the variable format specifications. There are NS integers indicating the levels of the factors, followed by the response vector of length NYR. The factor level codes must be consecutive in-
tegers beginning with 1 . If the codes are not consecutive integers a user transformation subroutine (TRANS) can be written to put the codes in the proper form. The response vector can also be transformed by this subroutine. If the data is on cards and if N is the number of cards per observation, N blank cards must follow last data card.

## 5. Axis Title Card <br> Cols.

1-28 YAXIS-Title for profile lines.
29-56 XAXIS-Title for factor on horizontal axis.
57-80 REPEAT-Title for factor given control value 116 on card 7.
6. General Title

Cols.
1-79 Title—general title for profile
800 or blank $=$ further plots involving present data follow
$1 \quad=$ last plot for this data set (needed only if further data sets follow).

## 7. Factor Control Card

On this card the user specifies the profiles he wants plotted by assigning plotting codes to the factors in the experiment. One of the codes that appear below must be assigned to the response and one to each factor to indicate how each is used for the profiles.

The plotting codes are
$I=\begin{aligned} & \text { Profiles for } I^{\text {th }} \\ & \text { sponse) }\end{aligned}$
113 = Profiles combined over levels of factor (or response).
$114=$ This factor (or response) used for vertical axis.
$115=$ This factor (or response) used for horizontal axis.
$116=$ Separate profiles plotted for each level of this factor (or response).

Codes 114, 115, and 116 can appear only once for each set of profiles. The NS +1 codes must be punched on the card in the following format. The remaining fields must be blank.

Cols.
1-5 Plotting Code for response vector.
6-10 Plotting Code for factor 1.
11-15 Plotting Code for factor 2.

56-60 Plotting Code for factor 10.
80 NES - horizontal scale indicator
NES $=0$ horizontal scale equally spaced.
$=1$ horizontal scale unequally spaced; scale card follows:
8. Horizontal Scale Card (optional, needed only if column 80 of card 7 is 1 )

Cols.
1-5 Scale position for level 1
6-10 Scale position for level 2 as many fields of 5 as needed.

Repeat cards 5, 6, 7, and 8 if needed for each additional set of profiles. Repeat cards 1 to 8 for each additional data set.

## DATA TRANSFORMATION ROUTINE

The option for the user to transform the data with a FORTRAN subroutine is provided. The placement of the routine depends upon the system that is being used. The IBM 360 or 370 series requires that this routine immediately precede the object deck.

Suppose in the example the three levels of phosphorus (factor C) are coded 0,2 , and 5. These codes need to be changed to 1,2 , and 3 respectively. Suppose also that the user wants to examine the response on the square root scale. The transformation routine could be,

```
    SUBROUTINE TRANS (Y, I, N, DS)
    REAL Y(N)
    INTEGER DS, I(11)
    Y(2) = SQRT (Y(1))
    IF (I(4)-2) 1, 2, 3
1 I(4)=1
    GO TO 4
2 I(4)=2
    GO TO 4
3 I(4)=3
4 \text { RETURN}
END
```

All arguments of the subroutine are trans-
ferred from the calling routine and are not specified by the user. The arguments N and DS should be explained further. N is the length of the response vector, DS is the number of the data set that is being read. If a job contains more than one data set for which different transformations are required, an IF statement must be inserted directing program flow to the appropriate transformations for each data set.

Notice that although the levels of the third factor are being recoded, the subscript for I is 4 . If the $\mathrm{j}^{\text {th }}$ factor is being recoded the subscript of I must be $j+1$. This is necessary since the response is considered the first factor by the program.

## PROGRAMMING CONSIDERATIONS

Definitions:

1. $\mathrm{NS}=$ Number of factors.
2. $\operatorname{MAX}(j+1)=$ Highest level of $\mathrm{j}^{\text {th }}$ factor.
MAX (1) $=$ Length of response vector. $\mathrm{NS}+1$
3. $\operatorname{MXSTR}=\Pi \quad \operatorname{MAX}(\mathrm{j})$
4. $\mathrm{NF}=$ Number of format cards.
5. JMAX $=$ Difference between largest and smallest relative levels of horizontal axis factor.

## Restrictions:

1. $\mathrm{NS} \leq 10$
2. $\operatorname{MAX}(\mathrm{j}) \leq 99$ all factors not used on vertical axis. $\leq 12$ any factors used on vertical axis.
3. $\mathrm{NF} \leq 3$
4. $1 \leq$ JMAX $\leq 110$
5. $\mathrm{MXSTR} \leq 5000$

If a problem is encountered where MXSTR $>5000$ the following changes should be made in the main program:

1. Set MSTR $=$ MXSTR .
2. Change dimension of Y and IC to Y (MXSTR) and IC (MXSTR).

## DESCRIPTION OF OUTPUT

Information from the first control card is printed on the first page of output. Following this is the variable format for this data set and a listing of the factor means.

For each job concerning this data set the actual profiles are preceded by a page containing means to be used on the profiles and the control values for each factor.

A profile for each level of the factor or response selected for the vertical axis is then plotted at each level of the horizontal axis. Numbers or letters are used as plotting characters to represent the levels of the vertical axis factor; levels 1 to 9 are represented by the integers 1 to 9 , level 10 by 0 , level 11 by A, and level 12 by B. A line segment consisting of the appropriate plotting character joins each point along the horizontal axis.

Overplots occur if the mean values of two or more profiles at any point on the horizontal axis are the same. These are listed on the page following the profiles in the form $\mathrm{Z}(\mathrm{I}, \mathrm{J})=\mathrm{M}$, where $I$ is the level of the vertical axis, $J$ is the point on the horizontal axis, and $M$ is the mean value.

## SAMPLE PROBLEM

The sample problem involves the same factors as discussed. The response vector is seedling height growth measured $1,2,4$, and 8 years after application of the fertilizer. The data cards have a 1 punched in the field for the first factor. This dummy factor is included in order that a profile for any one of the factors alone can be plotted.

The first set of profiles uses the dummy factor for the vertical axis and plots the profiles for phosphorus at time 1, summing over the other factors. The second set of profiles plots the means for the two levels of $\mathrm{N}(1=\mathrm{N}$ present, $2=\mathrm{N}$ absent) for each level of phosphorus at time 1 , while the third and fourth profiles represent the nitrogen by phosphorus profiles for each species.

Profile 1 shows that the mean height growth for the trees treated with the 3rd level of phosphorus is greater than the growth of the trees treated with levels 1 and 2. Profile 2 demonstrates the nitrogen by phosphorus in-
teraction. The mean height growth for phosphorus levels 2 and 3 is greater when nitrogen is present, while the mean growth for phosphorus level 1 is greater when nitrogen is absent.

Profiles 3 and 4 illustrate the lack of threeway interaction. The nitrogen by phosphorus profiles for each species are parallel to each other. Finally, profile 5 demonstrates the use of an additional control card to read relative values of the horizontal axis. The heights were measured at times 1, 2, 4, and 8. Therefore
the horizontal axis distance between measurements 2 and 3 is twice the distance between times 1 and 2. This set of profiles also shows the increase in mean height growth with time, and the fact that the species 2 means are greater than the means of species 1 for each time period. Also, the difference is of the same magnitude in each time period.

## Literature Cited

Winer, B. J.
1962. Statistical principles in experimental design. New York: McGraw Hill.

## ACKNOWLEDGMENT

The author wishes to thank Dr. Donald Seegrist, Biological Statistician at the Northeastern Forest Experiment Station for his encouragement and suggestions for improvement during the development of the program.

## CONTROL DECK FOR SAMPLE PROBLEM



## OUTPUT FOR SAMPLE PROBLEM

```
PROFILE ANALYSIS, CGNTROL INFCRMATICN FOR THIS DATA SET
NO. SUBSCRIPTS(FACTORS)=
NO. Y'S READ IN =
NO. FORMAT CARDS =
PRINT MEANS =
DATA FILE NO. =
COUNT GPTION, ICNT =
NO. CELLS NOT INCLUDED IN DESIGN FOR COUNT OPTION, NDES= 0
MAXIMUM VALUE OF EACH SUBSCRIPT, FIRST SUBSCRIPT REFERS TO LENGTH OF Y VECTOR ON DATA CARD
```



```
FORMAT OF DATA
(4I2,7X,4F5.0)
TOTAL NO. OF RECORDS ON DATA SET = 16
FACTOR SUBSCRIPTS, FIRST IS NO. OF LEVELS OF OBSERVATION VECTOR
```


profile analysis, means of profile cell plots for job 1


PROFile ANALYSIS, MEANS OF PROFILE CELL PLOTS FOR JOB 2
FACTOR LEVELS $\quad 0$ INDICATES THE LEVELS OF THIS FACTOR ARE TO BE THE LEVELS OF THE X AXIS INDICATES FACTOR FOR HORIZONTAL AXIS
INDICATES FACTOR FOR VERTICAL AXIS $n$
$\stackrel{a}{-}$
$a$
CONTROL VALUES

$$
\begin{aligned}
& \mathrm{X} \\
& \mathrm{Y} \text { INDICATES }
\end{aligned}
$$

EACH OF ITS LEVELS. A SEPARATE PROFILE

CONTRQL VALUES FOR FACTOR SUBSCRIPTS
113 IMPLIES CELL MEANS ARE SUMMED OVER THIS SUBSCRIPT. 113 IMPLIES CELL MEANS
114 LEVELS OF Y AXIS.
115
LEVELS OF X AXIS.
115 LEVELS OF X AXIS.
116 PRODUCE A SEPARATE PROFILE PLOT FQR EACH LEVEL OF THIS SUBSCRIPT
$1-111$ THIS LEVEL OF LTH FACTOR IS TO BE PLOTTED. $\propto$

profile analysis, job= 2 nitrogen x phosphorus combined species time a
$\mathbf{Y}=$ NITROGEN
$* * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * * *$

| 100.000 |
| :--- |
| 80.0000 |
| 60.0000 |
| 40.0000 |
| 20.0000 |
| .0 |






1111
**


2222
$1111^{111}$
$*$
$*$
$*$
$*$

1111
111
1111
1111

PROFILE ANALYSIS, MEANS OF PROFILE CELL PLOTS FOR JOB 3
FACTOR LEVELS 0 INDICATES THE LEVELS OF

$$
\left.\begin{array}{lllllllllllllll}
1 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & & & & \\
& & & & & & & & & Y & -Y 12 & \text { MEAN } \\
& 1 & 1 & 2 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & & & & \\
\text { NOBS }
\end{array}\right)
$$

0 INDICATES THE LEVELS OF THIS FACTOR ARE TO BE THE LEVELS OF THE X AXIS

* INDICATES THE FACTOR FOR WHICH A SEPARATE PROFILE PLOT IS PRODUCED FOR
INDICATES FACTOR FOR HORIZONTAL AXIS. $x$ INDICATES FACTOR FOR HORIZONTAL AXIS
y INDICATES FACTOR FOR VERTICAL AXIS

$$
\text { LEVELS OF } X \text { AXIS }
$$

$$
\begin{array}{ccr}
\text { LEVELS OF X AXIS } & \\
1 & 2 & 3 \\
\begin{array}{rrr}
0.0 & 30.00 & 60.00 \\
1 & 1 & 1 \\
40.00 & 10.00 & 40.00 \\
1 & 1 & 1
\end{array}
\end{array}
$$

CONTROL VALUES FOR FACTOR SLBSCRIPTS
113 IMPLIES CELL MEANS ARE SUMMED OVER THIS SUBSCRIPT.
114 LEVELS OF Y AXIS.
115 LEVELS OF $x$ AXIS.
116 PRODUCE A SEPARATE PROFILE PLOT FOR EACH LEVEL OF THIS SUBSCRIPT
$+4$
n $\pm$
$\sim 』$
$\rightarrow$
$\simeq$
control values
80.0000
60.0000

LEVELS OF $X=$ PHOSPHORUS
profile analysis, means of profile cell plots for job 3


profile analysis, job= 3 nitrogen $X$ phosphorus each species time 1
$\underset{* * * * * * * * * * * * *}{r}$
*********************************************

profile analysis, means of profile cell plots for job


PROFILE ANALYSIS, JOB= 4 SPECIES X TIME


