Eutamias minimus and E. amoenus : morphological cluster analysis

Sandra Elaine Anderson
Portland State University

Let us know how access to this document benefits you.
Follow this and additional works at: https://pdxscholar.library.pdx.edu/open_access_etds
Part of the Animal Sciences Commons, and the Biochemistry, Biophysics, and Structural Biology Commons

Recommended Citation

10.15760/etd.2262

This Thesis is brought to you for free and open access. It has been accepted for inclusion in Dissertations and Theses by an authorized administrator of PDXScholar. For more information, please contact pdxscholar@pdx.edu.
Cluster analysis of a large body of data on 180 Oregon specimens of *Eutamias minimus* and *E. amoenus* suggest that overall length of skull, basal length of skull and length of palate are taxonomically critical. If their sum is less than 71.3 millimeters the animal is *E. minimus*. If the sum is greater than 72.7 millimeters the animal is *E. amoenus*. If the sum is between 71.3 and
72.7 millimeters, other factors must be considered before the animal can be identified. Of the 180 specimens, there were 60 *E. minimus*, 114 *E. amoenus*, 2 hybrids, 2 unidentifiable and 2 mismatched skulls and skins.
EUTAMIAS MINIMUS AND E. AMOENUS:
MORPHOLOGICAL CLUSTER
ANALYSIS

by

SANDRA ELAINE ANDERSON

A thesis submitted in partial fulfillment of the requirements for the degree of

MASTER OF ARTS
in
BIOLOGY

Portland State University
1974
TO THE OFFICE OF GRADUATE STUDIES AND RESEARCH:

The members of the Committee approve the thesis of Sandra Elaine Anderson presented August 5, 1974.

Richard B. Forbes, Chairman

Richard S. Petersen

Quentin D. Clarkson

APPROVED:

Earl Fisher, Jr., Head, Department of Biology

David T. Clark, Dean of Graduate Studies and Research

August 5, 1974
ACKNOWLEDGMENTS

The author would like to thank R. Forbes, B. Vertes, R. Storm, A. Walker, J. Kirk, L. Turner, D. Breakey and the American Museum of Natural History for providing study specimens. D. Clarkson loaned an invaluable reference work. D. Shell helped debug the computer program. The Department of Biology at Portland State University paid for computer time. Particular thanks go to R. Forbes for suggesting the problem and for kind encouragement. The paper would not have been completed without the typing skill of L. Alexander.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>v</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>vi</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>I INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>II MATERIALS AND METHODS</td>
<td>4</td>
</tr>
<tr>
<td>III THE ALGORITHM</td>
<td>15</td>
</tr>
<tr>
<td>IV RESULTS</td>
<td>22</td>
</tr>
<tr>
<td>V A SUGGESTED TAXONOMIC PROCEDURE</td>
<td>30</td>
</tr>
<tr>
<td>VI SEVEN PUZZLING CHIPMUNKS</td>
<td>36</td>
</tr>
<tr>
<td>LITERATURE CITED</td>
<td>40</td>
</tr>
<tr>
<td>APPENDIX</td>
<td>41</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>TABLE</th>
<th>DESCRIPTION</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>I</td>
<td>Sample Size, Range and Mode of Rejected Variables</td>
<td>11</td>
</tr>
<tr>
<td>II</td>
<td>Distance Matrix at 10-Cluster Stage</td>
<td>23</td>
</tr>
<tr>
<td>III</td>
<td>Information Concerning the Frequency Distributions of Size Variables</td>
<td>25</td>
</tr>
<tr>
<td>IV</td>
<td>Distributions of Color Variables</td>
<td>28</td>
</tr>
<tr>
<td>V</td>
<td>Pertinent Information on a Series of Chipmunks Taken Near Madras</td>
<td>38</td>
</tr>
<tr>
<td>VI</td>
<td>Pertinent Information on a Series of Chipmunks Taken at Skeleton Cave</td>
<td>39</td>
</tr>
</tbody>
</table>
CHAPTER I

INTRODUCTION

It is commonly believed that two species of small chipmunks, called *Eutamias minimus* and *E. amoenus*, inhabit eastern Oregon. Some animals are easily assigned to one species or the other, but all too often the biologist is confronted with an animal he cannot identify. Dr. Richard B. Forbes, an expert on chipmunks, had been frustrated for many years by a set of specimens in the Portland State University collection, and it was he who suggested this study.

Ingles (1965) gives three rules for identifying chipmunks. 1) If the skull length is greater than 31.5 mm. and the animal was living in timber, it is *E. amoenus*. If the skull length is less than 31.5 mm. and the animal was living in sagebrush, it is *E. minimus*. 2) If the skull is more than 31.0 mm. long, the animal is *E. amoenus*. 3) If the skull is less than 30.5 mm. long, the animal is *E. minimus*.

Hall and Kelson (1959) distinguish between the two species by bacular morphology. If the length of the bent tip of the baculum is less than 28 per cent of the shaft, the animal is *E. minimus*, otherwise it is *E. amoenus*. 
# LIST OF FIGURES

<table>
<thead>
<tr>
<th>FIGURE</th>
<th>PAGE</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 Dorsal view of cranium</td>
<td>6</td>
</tr>
<tr>
<td>2 Ventral view of cranium</td>
<td>7</td>
</tr>
<tr>
<td>3 Ventral view of half cranium</td>
<td>8</td>
</tr>
<tr>
<td>4 Side view of cranium</td>
<td>9</td>
</tr>
<tr>
<td>5 Mandible</td>
<td>10</td>
</tr>
<tr>
<td>6 BL X GL, ( x = \text{minimus}, o = \text{amoenus} )</td>
<td>31</td>
</tr>
<tr>
<td>7 BL X PL, ( x = \text{minimus}, o = \text{amoenus} )</td>
<td>32</td>
</tr>
<tr>
<td>8 GL X PL, ( x = \text{minimus}, o = \text{amoenus} )</td>
<td>33</td>
</tr>
</tbody>
</table>
The objection to this approach is that the small size of the baculum, about 3 mm., makes it difficult to prepare and easy to lose. Moreover, it is only present on half the animals. Hall and Keison go on to state "As yet no specific diagnosis has been framed that will serve to distinguish the two species in all parts of their geographic ranges. At any one place however, the two are distinguishable." They discuss differences in Nevada, California, Wyoming, and Montana, and refer the reader to Howell (1929) for other areas.

Howell (1929) emphasizes color. The Oregon race of *E. minimus*, *E. minimus pictus* in Howell's terminology, has "Head smoke gray, more or less mixed with light pinkish cinnamon...". An Oregon race of *E. amoenus*, *E. amoenus amoenus amoenus*, has "Top of head smoke gray, mixed with cinnamon...". The reader emerges from Howell more confused than ever.

Sutton and Nadler (1969) classified races of chipmunks by chromosome type. All subspecies of *E. amoenus* had Karyotype B. Some subspecies of *E. minimus* had Karyotype A, others, B. The race found in Oregon, *E. minimus scrutator*, had Karyotype B. Sutton and Nadler say "It is possible, but not likely, that the six subspecies of *E. minimus* with Karyotype B are erroneously classified and are referable to *E. amoenus*, which resembles
minimus closely in color, measurements, and bacular morphology."

Do the small eastern Oregon chipmunks form a single morphologically homogeneous group or do they form two morphologically distinct groups? If the latter, what is the distinction?

I sought to answer these questions by studying a large number of morphological features on a large number of chipmunks without prejudging their species. The chipmunks could then be classified by cluster analysis where, as Sokol (1966) says, "Assignment to the taxon is not on the basis of a single property but on the aggregate of the properties, and any pair of members of the class will not necessarily share every character." If the cluster analysis reveals the sample to be relatively homogeneous, then one could argue that the two species are not morphologically distinguishable in Oregon. If, on the other hand, the cluster analysis reveals the sample to be composed of two relatively distinct groups, then one could argue that the group containing the smaller chipmunks represents E. minimus and the other group E. amoenus. Examination of the two groups may suggest a simple classification rule that gives almost the same results as the cluster analysis. That rule could be used to identify any other small eastern Oregon chipmunk.
CHAPTER II

MATERIALS AND METHODS

One hundred and eighty post-juvenal individuals, including 97 males, 79 females, and 4 of undetermined sex, were examined. There were 47 from Portland State University (PSU), 45 from each of Oregon State University, Fish and Wildlife (OSUFW) and Oregon State University, Zoology (OSUZ), 17 from Alex Walker's private collection (AW), 8 from the Oregon State Board of Health (OSBH), 7 from the American Museum of Natural History (AMNH), 6 from the Malheur Environmental Field Station (MEFS), and 5 from Willamette University (WU). Their geographical origins were: Deschutes County: Cold Springs, 10; Skeleton Cave, 17; Middle Sister, 1; Black Butte, 2; Sisters, 2; Hay Creek, 1; McKenzie Pass, 1; other, 23. Grant County: 3. Harney County: Burns, 3; Malheur Refuge, 3; Steens Mountain, 1; other, 39. Hood River County: 2. Jackson County: 1. Jefferson County: Haystack Reservoir, 2; 10 miles S. of Madras, 2; 10 miles S. and 3 miles E. of Madras, 8; Warm Springs, 5. Klamath County: 4. Lake County: Fort Rock, 4; Warner Valley, 6; Warner Mountains, 5; Hart Mountain, 2; other, 7. Malheur County: Brogan, 1; Ironside, 7; other, 2. Umatilla County: 1. Wallowa County: Joseph, 1; other, 5. Wasco County:
The Dalles, 2. Wheeler County: Mitchell, 2; Shelton Park, 3.

The 34 skull measurements indicated in Figures 1-5 were selected as possible variables for analysis. Measurements were taken with a Helios dial caliper to the nearest 0.1 mm. Features present on both sides of the skull were measured on the left side whenever possible. The variable AM was rejected because its value depended on how thoroughly the skull had been cleaned. Two variables, TBS and AB, were rejected because of difficulty in proper positioning of the caliper tips. If a variable had too small a range, the information it could yield would not justify the effort of punching its values on cards. It was arbitrarily decided that a range of less than 2 mm. was too small. The 11 variables listed together with their sample sizes, ranges, and modes in Table I were so rejected.

Two small notches, one at the anterior tip of the rostrum and the other on the least interorbital constriction, occasionally occur. They were coded as RN and ION and scored as present or absent. They were excluded from the cluster analysis because there was no reasonable way to weight them, but they will be discussed in Chapter IV.

The body measurements total length and lengths of tail (T), hind foot (HF) and ear from notch (E) were taken from specimen tags. They had generally been recorded to the nearest 1 mm. The few ear and hind foot lengths
Figure 1. Dorsal view of cranium.
Figure 2. Ventral view of cranium.
Figure 3. Ventral view of half cranium.
Figure 5. Mandible.
TABLE I
SAMPLE SIZE, RANGE, AND MODE
OF REJECTED VARIABLES

<table>
<thead>
<tr>
<th>Variable</th>
<th>Sample Size</th>
<th>Range</th>
<th>Mode</th>
</tr>
</thead>
<tbody>
<tr>
<td>RWI</td>
<td>171</td>
<td>3.3-5.0</td>
<td>4.4</td>
</tr>
<tr>
<td>WI</td>
<td>172</td>
<td>2.7-4.3</td>
<td>3.5</td>
</tr>
<tr>
<td>IF</td>
<td>173</td>
<td>1.2-2.4</td>
<td>1.9</td>
</tr>
<tr>
<td>IFS</td>
<td>170</td>
<td>5.4-6.9</td>
<td>6.4</td>
</tr>
<tr>
<td>PW</td>
<td>170</td>
<td>3.9-5.5</td>
<td>4.7</td>
</tr>
<tr>
<td>WIN</td>
<td>171</td>
<td>1.6-2.7</td>
<td>2.2</td>
</tr>
<tr>
<td>TBL</td>
<td>172</td>
<td>7.2-9.0</td>
<td>8.0</td>
</tr>
<tr>
<td>OV</td>
<td>173</td>
<td>7.0-8.9</td>
<td>8.1</td>
</tr>
<tr>
<td>CHFM</td>
<td>170</td>
<td>8.5-10.3</td>
<td>9.6</td>
</tr>
<tr>
<td>LD</td>
<td>172</td>
<td>3.1-5.0</td>
<td>4.3</td>
</tr>
<tr>
<td>AT</td>
<td>170</td>
<td>3.8-5.7</td>
<td>4.6</td>
</tr>
</tbody>
</table>
recorded to 0.5 mm. were haphazardly rounded up or down to a whole millimeter. Tail length was subtracted from total length to obtain head and body length (HBL). The four variables HBL, T, HF, and E were analyzed.

Chipmunks often lose part of their tails to the slings and arrows of outrageous fortune. The following procedure was used to determine which tails had been accidentally shortened. The modal value for T in a sample of 180 chipmunks was 85 mm. The largest value was 112 mm., and 112 - 85 = 27. If we assume a symmetrical distribution, which is reasonable because the curve was a beautiful bell, then the smallest reasonable value for T is 85 - 27 or 58. The 8 values less than 58 mm. were rejected. One other value of T, that of the OSUZ specimen CM 219, was rejected. Total length was 201 mm., tail length was 70 mm., so HBL = 131 mm., the largest value that variable attained. The skull was only 29.9 mm. long. Such a chipmunk would have been grotesque. I believe there was an error on the specimen tag. I deleted both HBL and T from the data on that individual.

Two color variables, the color of the tail underside (TU) and the body side (S) were analyzed. Four skins (PSU 194, 358, 1135, and 1223) were chosen as standards for TU. Color varied from a sharply margined broad band of deep rich brownish orange on PSU 194 to a narrow indistinct band of pale yellow mixed with black on PSU 1223. Four
other skins (PSU 872, 876, 1217 and 1224) were chosen as standards for S. PSU 872 was bright orange from shoulder to hip. PSU 1224 had a touch of pale orange at the base of the ribcage. PSU 876 and 1217 were intermediate. Every other skin was assigned a value of 1, 2, 3 or 4 for each variable, according to which of the standards it most resembled. The OSUFW skins were viewed under fluorescent light, WU skins under incandescent light, and the others under diffused daylight. It is possible that slightly different values would have been obtained if all the skins had been viewed under the same light. The value of S was estimated on molting animals above the molt line if and only if the line was more than half way down the body. Color contrast between summer and winter pelage on some molting animals was striking, summer pelage being much brighter.

Sex, place of capture, habitat and putative species were noted but not analyzed. A few individuals from the OSUFW collection taken in late fall were reproductively immature. Only individuals with fully erupted third molars were examined.

Data were punched onto IBM cards and fed into a CDC3300 computer. Skull variables were punched as tenths of millimeters and body variables as millimeters. Color variables were multiplied by 15 so that the difference
between the largest and smallest values was 15(4) - 15(1) = 45, or about the same as the difference between the largest and smallest values of GL (343 - 287 = 56) and HBL (131 - 94 = 37). Missing values were coded as -10. Eighty-nine chipmunks were missing from one to five values for a total of 155 missing values out of a theoretical 26 values per chipmunk X 180 chipmunks = 4,680 values.
CHAPTER III

THE ALGORITHM

The computer program is listed in the Appendix. It was simplified and modified from one given by Anderberg (1973). Chipmunks are plotted as points in a multidimensional space, a measure of distance is calculated between every pair of points, and points that are close to each other are clustered together.

Phenetic distances are calculated in lines 26-42. Euclidean distance, normally the first choice for a distance, is not suitable. Suppose two variables had been measured on three chipmunks to give the triples (1, 5, 7) and (2, 4, 8). The Euclidean distance between the two is the square root of the sum of the squares of the differences of the co-ordinates, in this case $\sqrt{3} = 1.7$. Now suppose the third variable had not been measured on the first chipmunk, perhaps because of a broken bone, so that the triples are (1, 5, -10) and (2, 4, 8). If we ignore the missing value, so that the chipmunks are compared in two dimensions instead of three, the distance decreases to $\sqrt{2} = 1.4$. Ignoring the missing variable is equivalent to the obviously unwarranted assumption that the two chipmunks had the same value for that variable.
The Euclidean distance must be modified so that distances calculated in spaces of different dimensions are comparable. In particular, we would like the diameter of the unit cube (the distance between the point (0, 0, ...0) and the point (1, 1, ...1)) to be independent of the number of dimensions. This can be done by dividing the Euclidean distance by the Euclidean diameter of the unit cube, so that the unit cube has phenetic diameter 1. The unit cube in p-space has Euclidean diameter $\sqrt{p}$. The phenetic distance between Chipmunk A and Chipmunk B is defined to be

$$\sqrt{\frac{\sum (a_i - b_i)^2}{p}}$$

where $a_i$ is the value of variable $i$ on Chipmunk A, $b_i$ is its value on Chipmunk B, the summation is over all variables evaluated on both chipmunks, and $p$ is the number of terms in the sum.

Phenetic distance is neither a metric nor a pseudometric. It is less than the largest difference between values and greater than the mean of the differences.
The phenetic distance between the points (1, 5, 7) and (2, 4, 8) considered above is 1. The phenetic distance between (1, 5, -10) and (2, 4, 8) is also 1.

In order to save core storage, the real values for phenetic distance were converted to integers. An integer takes only half as much core storage as a real number. In this case phenetic distances were less than 56, or well below $2^6 = 64$. The distances were multiplied by $2^9 = 512$ (equivalent to moving the binary point three places to the right) and rounded off, resulting in integers less than $2^{15} = 32,768$. Distances were numbered consecutively from 1 to 16,110.

Function LFIND (lines 1-11) retrieves the distances. Consider the array

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>4</td>
<td>5</td>
<td>6</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>10</td>
</tr>
</tbody>
</table>

Numbers across the top and down the side represent chipmunks. Numbers in the body are identification numbers of distances. If we desire the distance between Chipmunk 2 and Chipmunk 4, LFIND tells us to look at the 5th distance.

The problem of selecting a classification method is
non-trivial. I looked for a method that involved only a few simple calculations. Two methods, single linkage and maximum linkage, are sufficiently simple. Anderberg (1973) states that single linkage does not work well if clusters are poorly defined. Since I believed my clusters were poorly defined, I chose maximum linkage. This means the distance between Clusters A and B is the largest of all distances between Chipmunk a in A and Chipmunk b in B, i.e. the diameter of A union B.

To see how maximum linkage works, consider the five points

\[ x_5 \quad x_4 \quad x_2 \quad x_1 \quad x_3 \]

The picture can be described by the matrix

\[
\begin{array}{cccc}
1 & 2 & 3 & 4 \\
2 & 3 & & \\
3 & 2 & 5 & \\
4 & & 1 & 6 \\
5 & & 7 & 9 & 3
\end{array}
\]

Initially each individual is regarded as a cluster with one member. In lines 52-62 the machine finds the smallest entry in each row, and in lines 64-71 it finds the smallest entry in the entire array, in this case 1. Clusters 2 and 4 are merged in lines 78-82 and will henceforth be called Cluster 2. Cluster 5 is now the 4th
cluster in the list. Lines 84-88 permit the machine to keep track of the original identification numbers.

The matrix is now

\[
\begin{array}{ccc}
1 & 2 & 3 \\
2 & & \\
3 & 2 & \\
5 & 7 & 9 \\
\end{array}
\]

We must now decide what distance Cluster 2 should be from all the other clusters. With maximum linkage, the distance is as large as possible. To find the distance from Cluster 1 to new Cluster 2, we compare the distance from Cluster 1 to old Cluster 2 (3 units) with that to old Cluster 4 (4 units). The larger of these is 4, so the distance from Cluster 1 to new Cluster 2 is 4. The distances from Clusters 3 and 5 to new Cluster 2 are similarly defined, yielding the matrix

\[
\begin{array}{ccc}
1 & 2 & 3 \\
2 & 4 & \\
3 & 2 & 6 \\
5 & 7 & 4 & 9 \\
\end{array}
\]

This brings us to line 100. In lines 104-124 we again look for the smallest entry in each row. If it is not time for output, we go back to line 64 and repeat the procedure. Next merge Clusters 1 and 3 and update
the matrix:

\[
\begin{array}{c|cc}
  & 1 & 2 \\
\hline
1 & 1 & 2 \\
2 & 6 & 2 \\
3 & 9 & 4 \\
\end{array}
\]

and once more: merge Clusters 2 and 5 to obtain

\[
\begin{array}{c|c}
  & 1 \\
\hline
1 & 1 \\
2 & 9 \\
\end{array}
\]

There are now only two clusters so it is time to stop. Lines 126-148 direct the machine to print the results. In this case, Cluster 1 contains Chipmunks 1 and 3 and Cluster 2 contains Chipmunks 2, 4 and 5. Note that even though Chipmunk 2 resembles Chipmunk 1 more than Chipmunk 5, 2 and 5 were put in the same final cluster, and 2 and 1 in different clusters.

A disadvantage of maximum linkage is that results can depend on the order in which numbers are assigned to individuals. Consider the three points

\[
\begin{array}{c|ccc}
  & x & x & x \\
\hline
1 & 1 & 2 & 3 \\
\end{array}
\]

The distance matrix is

\[
\begin{array}{c|cc}
  & 1 & 2 \\
\hline
1 & 1 & 2 \\
2 & 1 & 2 \\
3 & 2 & 1 \\
\end{array}
\]
There is no unique smallest entry. Somehow we must decide which of the two l's should be called the smallest. The algorithm always selects the last appearance of a minimum element as the "smallest" value. In this case the point in the middle is merged with the point on the right. However, if the points are renumbered

\[ X \quad X \quad X \]

\[ 3 \quad 2 \quad 1 \]

the distance matrix is the same but now the point in the middle is merged with the point on the left. Assignment of intermediate individuals to clusters can be altered by shuffling the data deck.
CHAPTER IV

RESULTS

The final two clusters were a set of 62 smaller chipmunks, henceforth called the minimus group, and a set of 118 larger chipmunks, henceforth called the amoenus group. The distance matrix at the 10-cluster stage is displayed in Table II. Distances are rounded to two significant figures and the final two zeros dropped. Clusters 27 and 69 form a bridge between the amoenus clusters 11, 20, 104 and 18 and the minimus clusters 11, 20, 104 and 18. They were ultimately placed in the amoenus group.

Cluster 69, the smallest cluster, contained 5 chipmunks. One of these, Chipmunk 114 (OSUZ 1518) had a value of only 30.4 for greatest skull length (GL). The next smallest value for GL in the amoenus group was 30.9 which is much larger. Chipmunk 114 seemed suspect, so its raw data were checked. It had a small skull and orange color and according to the specimen tag came from a fir-juniper area (E. amoenus habitat). Chipmunk 117 (OSUZ 1531) had a large skull and yellow and gray colors and came from an aspen-sage area (E. minimus habitat). Both were taken in the Malheur refuge. I believe the
TABLE II

DISTANCE MATRIX AT 10-CLUSTER STAGE

<table>
<thead>
<tr>
<th></th>
<th>8</th>
<th>4</th>
<th>1</th>
<th>2</th>
<th>27</th>
<th>69</th>
<th>11</th>
<th>20</th>
<th>104</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>62</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>71</td>
<td>62</td>
<td>58</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>78</td>
<td>61</td>
<td>70</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>69</td>
<td>86</td>
<td>69</td>
<td>63</td>
<td>66</td>
<td>55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>102</td>
<td>91</td>
<td>87</td>
<td>75</td>
<td>71</td>
<td>62</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td>122</td>
<td>103</td>
<td>98</td>
<td>94</td>
<td>83</td>
<td>64</td>
<td>54</td>
<td></td>
<td></td>
</tr>
<tr>
<td>104</td>
<td>135</td>
<td>115</td>
<td>110</td>
<td>110</td>
<td>95</td>
<td>74</td>
<td>72</td>
<td>62</td>
<td></td>
</tr>
<tr>
<td>18</td>
<td>147</td>
<td>130</td>
<td>127</td>
<td>117</td>
<td>108</td>
<td>91</td>
<td>76</td>
<td>70</td>
<td>65</td>
</tr>
</tbody>
</table>
skulls and skins (with specimen tags) had been interchanged, so that Chipmunk 114 was an *E. minimus* skull paired with an *E. amoenus* skin and tag, and Chipmunk 117 was an *E. amoenus* skull paired with an *E. minimus* skin and tag. These two specimens were omitted from further analysis.

Contrary to expectations, series of specimens taken at the same time and place were not clustered together at the early stages. The AMNH series from Ironside, Malheur County, was divided between Clusters 18 and 20 (although 4 of them were clustered together at the 100-cluster stage). The AW Warner Valley series was scattered among Clusters 18, 104 and 20. AW Warner Mountains animals went into Clusters 2, 27 and 4. PSU Cold Springs chipmunks were placed in Clusters 2, 4 and 8. One of the PSU Madras chipmunks was put in Cluster 18 and the others in Cluster 11. Most interestingly, the PSU Skeleton Cave series had some individuals in the minimus group and others in the amoenus group.

Chipmunk 50 (OSUFW 981) was not clustered with any other chipmunk until the 10-cluster stage. It did not appear to be particularly unusual. All other chipmunks were clustered at the 20-cluster stage or earlier.

Table III contains information on the frequency distributions of the size variables. Columns 1-8 are the sample size, range, median and mode for the minimus and
<table>
<thead>
<tr>
<th>Var.</th>
<th>Minimus Group</th>
<th>Amoenus Group</th>
<th>Overlap</th>
</tr>
</thead>
<tbody>
<tr>
<td>BL</td>
<td>54</td>
<td>23.4-26.0</td>
<td>24.9</td>
</tr>
<tr>
<td>GL</td>
<td>61</td>
<td>28.7-31.1</td>
<td>29.7</td>
</tr>
<tr>
<td>PL</td>
<td>55</td>
<td>13.2-15.2</td>
<td>14.4</td>
</tr>
<tr>
<td>OR</td>
<td>62</td>
<td>9.4-10.8</td>
<td>10.2</td>
</tr>
<tr>
<td>CMa</td>
<td>62</td>
<td>14.1-16.8</td>
<td>15.4</td>
</tr>
<tr>
<td>ON</td>
<td>61</td>
<td>9.8-11.4</td>
<td>10.5</td>
</tr>
<tr>
<td>PM</td>
<td>62</td>
<td>6.0-7.2</td>
<td>6.6</td>
</tr>
<tr>
<td>LIC</td>
<td>62</td>
<td>6.1-7.1</td>
<td>6.6</td>
</tr>
<tr>
<td>CHBO</td>
<td>61</td>
<td>9.8-11.0</td>
<td>10.3</td>
</tr>
<tr>
<td>ZB</td>
<td>62</td>
<td>15.4-17.9</td>
<td>16.7</td>
</tr>
<tr>
<td>UD</td>
<td>62</td>
<td>6.6-8.2</td>
<td>7.2</td>
</tr>
<tr>
<td>CBZ</td>
<td>62</td>
<td>14.3-16.0</td>
<td>15.2</td>
</tr>
<tr>
<td>Var.</td>
<td>Minimus Group</td>
<td>Amoenus Group</td>
<td>Overlap</td>
</tr>
<tr>
<td>------</td>
<td>---------------</td>
<td>---------------</td>
<td>----------</td>
</tr>
<tr>
<td>N</td>
<td>61</td>
<td>8.0-9.6</td>
<td>9.1</td>
</tr>
<tr>
<td>CT</td>
<td>62</td>
<td>4.1-5.8</td>
<td>4.8</td>
</tr>
<tr>
<td>AA</td>
<td>62</td>
<td>11.7-14.5</td>
<td>13.3</td>
</tr>
<tr>
<td>RWZ</td>
<td>62</td>
<td>8.2-10.2</td>
<td>9.2</td>
</tr>
<tr>
<td>CBPO</td>
<td>61</td>
<td>9.5-11.0</td>
<td>10.2</td>
</tr>
<tr>
<td>CRA</td>
<td>60</td>
<td>12.1-14.8</td>
<td>13.7</td>
</tr>
<tr>
<td>RWM</td>
<td>61</td>
<td>6.3-8.5</td>
<td>7.4</td>
</tr>
<tr>
<td>CHP</td>
<td>62</td>
<td>7.1-9.4</td>
<td>7.9</td>
</tr>
<tr>
<td>HBL</td>
<td>61</td>
<td>92-116</td>
<td>104</td>
</tr>
<tr>
<td>T</td>
<td>58</td>
<td>60-97</td>
<td>84</td>
</tr>
<tr>
<td>HF</td>
<td>59</td>
<td>24-31</td>
<td>29</td>
</tr>
<tr>
<td>E</td>
<td>44</td>
<td>10-17</td>
<td>15</td>
</tr>
</tbody>
</table>

* See text.
amoenus groups. Column 9 is the overlap between the two groups. An isolated extreme value might indicate a misclassified individual, an abnormal individual, or a measurement error. If an extreme value was 2 units or more from the next value in the same group, it was not considered when calculating the overlap. Isolated extreme values were ignored in the 14 overlaps marked with a star.

Since BL, GL and PL had the shortest overlaps, they should be the most useful taxonomically. There were 2, 3 and 4 chipmunks in the overlaps of BL, GL and PL. The variables OR, CNA and ON had 26, 28 and 33 chipmunks in their overlaps and are therefore much less useful. All of the chipmunks that were intermediate on BL, GL or PL were also intermediate on OR, CNA and PL. Since all the other variables had overlaps including either the median or mode of one or both groups either as an end point or an interior point, they are taxonomically useless.

Table IV shows the distribution of the color variables over the two groups. The amoenus group tends to be more orange, particularly on the tail underside, but intermediate colors are well represented in both groups. Color does not seem to be a reliable guide to identification.

The variables ION and RN were tallied for both groups. The interorbital notch was present on 25 of 62 minimus members and 41 of 116 amoenus members, that is, it
# TABLE IV
## DISTRIBUTION OF COLOR VARIABLES

<table>
<thead>
<tr>
<th>Color TU</th>
<th>Minimus Group</th>
<th>Amoenus Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1 (orange)</td>
<td>0</td>
<td>66</td>
</tr>
<tr>
<td>2</td>
<td>6</td>
<td>44</td>
</tr>
<tr>
<td>3</td>
<td>39</td>
<td>6</td>
</tr>
<tr>
<td>4 (yellow)</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1 (orange)</td>
<td>2</td>
<td>51</td>
</tr>
<tr>
<td>2</td>
<td>15</td>
<td>53</td>
</tr>
<tr>
<td>3</td>
<td>28</td>
<td>12</td>
</tr>
<tr>
<td>4 (gray)</td>
<td>17</td>
<td>0</td>
</tr>
</tbody>
</table>
occurred about a third of the time in both groups. The rostral notch was present on 54 of 116 amoenus members and on only 4 of 61 minimus members. It seems that *E. amoenus* is much more likely to have a rostral notch than is *E. minimus*. Presence or absence of a rostral notch is not diagnostic for a single individual, but knowledge of its frequency in a population may be useful.

As expected, chipmunks in the minimus group did not come from heavy timber and those in the amoenus group generally did not come from pure sagebrush.
CHAPTER V

A SUGGESTED TAXONOMIC PROCEDURE

Each chipmunk can be identified with a point in a 3-dimensional space with co-ordinates BL, GL and PL. Figures 6, 7 and 8 are projections of the set of chipmunk-points on the planes PL = constant, GL = constant and BL = constant. If the constants are zero, the projections are on the co-ordinate planes. The set is hourglass-shaped with its long axis making an angle of about 45 degrees with each of the co-ordinate axes. The minimus group is in the lower bowl, the amoenus group in the upper. The two groups intermingle in the neck. The groups can be separated by a plane passing through the neck of the hourglass and perpendicular to its long axis. We seek its equation.

An arbitrary plane has equation

\[ kBL + mGL + nPL = ka + mb + nc \]

where \((a,b,c)\) is any point on the plane and \((k,m,n)\) are direction numbers for its normal.

The long axis of the hourglass is normal to the separating plane. It has direction numbers \((1,1,1)\). The equation we seek is of the form

\[ BL + GL + PL = a + b + c. \]
Figure 6. BL X GL, x = minimus, o = amoenus.
Figure 7. BL x PL, x = minimus, o = amoenus.
Figure 8. GL X PL, x = minimus, o = amoenus.
Choose \((a,b,c)\) to be the point of intersection of the separating plane with the long axis of the hourglass. The intersection of the separating plane with the horizontal plane \(PL = c\) through \((a,b,c)\) is a line whose equations are

\[
BL + GL + PL = a + b + c \quad (1)
\]

\[
PL = c.
\]

Consider Figure 6 as a projection of the hourglass on the plane \(PL = c\). The line (1) is approximately the solid line in Figure 6. Its equations are

\[
BL + GL = 56.9
\]

\[
PL = c.
\]

Thus \(a + b = 56.9\). \quad (2)

Similarly, by considering the vertical plane \(GL = b\) and Figure 7,

\[
a + c = 41.1. \quad (3)
\]

And, by considering the vertical plane \(BL = a\) and Figure 8,

\[
b + c = 46.0. \quad (4)
\]

The solution of equations (2), (3) and (4) is \(a = 26.0, b = 30.9\) and \(c = 15.1\). Their sum is 72.0. The equation of the separating plane is

\[
BL + GL + PL = 72.0.
\]

Prudence demands the separating plane be expanded to a thin slab. The slab should be thick enough to contain most of the hourglass neck but thin enough to miss the
bowls. My feeling is that a slab 1.4 units thick, extending 0.7 units above and below the separating plane as shown by the dotted lines in Figure 6, is neither too thin nor too thick. Taxonomists who are more ambiguity-tolerant and error-intolerant may prefer a thicker slab, perhaps as much as 2.0 units. More ambiguity-intolerant and error-tolerant taxonomists may prefer a thinner slab, but certainly no thinner than 0.6 units.

The three-dimensional chipmunk-space is partitioned into three regions. The region

\[ BL + GL + PL < 71.3 \]

represents \textit{E. minimus}. The region

\[ BL + GL + PL > 72.7 \]

represents \textit{E. amoenus}. The region

\[ 71.3 \leq BL + GL + PL \leq 72.7 \]

represents unidentifiable chipmunks.
CHAPTER VI
SEVEN PUZZLING CHIPMUNKS

Chipmunks that are unidentifiable by the method of Chapter V can sometimes be assigned to one species or the other by consideration of color, habitat, and other individuals in the same population.

Two chipmunks with intermediate values for GL were missing values for BL and PL. These were Chipmunk 57 (OSUFW 2198) with GL = 30.8 and Chipmunk 80 (OSUFW 1692) with GL = 31.1. Unfortunately they were intermediately colored and their habitats had not been recorded on the specimen tags. They will probably be forever unidentifiable.

Five chipmunks had BL + GL + PL between 71.3 and 72.7.

Chipmunk 109 (OSUZ 735) had BL = 25.9, GL = 30.9 and PL = 14.8 (sum = 71.4). It had a relatively orange tail underside and a bright orange side. It was taken in a yellow-pine area. On the basis of color and habitat, I am willing to call it E. amoenus.

Chipmunk 16 (PSU 1138) and Chipmunk 19 (PSU 1141) tended toward yellow and gray coloration. They were taken in a sagebrush area, as part of the PSU series from
10 miles south and 3 miles east of Madras. Table V contains pertinent information on the series. The only large gap is between Chipmunks 18 and 11. There is good reason to believe the series represents a single interbreeding population. They ought all belong to the same species. The mean of the sum of BL, GL and PL is well within the minimus region defined in Chapter V. If Chipmunk 18 is omitted, the mean of the sum is 70.9 which is still within the minimus region. The entire series should be assigned to the species *E. minimus*.

The two other chipmunks in the unidentifiable region are more interesting. Chipmunk 21 (PSU 1217) tended toward yellow and gray coloration. Chipmunk 27 tended toward orange. They were taken in a juniper area as part of the PSU series from Skeleton Cave. Table VI contains pertinent information on this series. There are large gaps between Chipmunks 28 and 21 and between Chipmunks 27 and 24. All the other gaps are small. All the animals were male, so sex differences cannot explain the discontinuities. I believe that Chipmunks 20 through 32 are *E. minimus*, Chipmunks 24 through 40 are *E. amoenus*, and Chipmunks 21 and 27 are hybrids.
TABLE V
PERTINENT INFORMATION ON
A SERIES OF CHIPMUNKS
TAKEN NEAR
MADRAS

<table>
<thead>
<tr>
<th>Chipmunk</th>
<th>BL</th>
<th>GL</th>
<th>PL</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>18</td>
<td>24.2</td>
<td>29.3</td>
<td>14.1</td>
<td>67.6</td>
</tr>
<tr>
<td>11</td>
<td>25.2</td>
<td>30.2</td>
<td>14.7</td>
<td>70.1</td>
</tr>
<tr>
<td>13</td>
<td>25.3</td>
<td>30.1</td>
<td>14.7</td>
<td>70.1</td>
</tr>
<tr>
<td>17</td>
<td>25.6</td>
<td>30.0</td>
<td>14.7</td>
<td>70.3</td>
</tr>
<tr>
<td>14</td>
<td>25.5</td>
<td>30.4</td>
<td>14.9</td>
<td>70.8</td>
</tr>
<tr>
<td>15</td>
<td>25.7</td>
<td>30.8</td>
<td>14.5</td>
<td>71.0</td>
</tr>
<tr>
<td>19</td>
<td>25.9</td>
<td>30.6</td>
<td>14.9</td>
<td>71.4</td>
</tr>
<tr>
<td>16</td>
<td>26.0</td>
<td>31.1</td>
<td>15.2</td>
<td>72.3</td>
</tr>
<tr>
<td>Mean</td>
<td>25.4</td>
<td>30.3</td>
<td>14.7</td>
<td>70.45</td>
</tr>
<tr>
<td>Chipmunk</td>
<td>BL</td>
<td>GL</td>
<td>PL</td>
<td>Sum</td>
</tr>
<tr>
<td>----------</td>
<td>----</td>
<td>----</td>
<td>----</td>
<td>------</td>
</tr>
<tr>
<td>20</td>
<td>24.9</td>
<td>29.5</td>
<td>14.2</td>
<td>68.6</td>
</tr>
<tr>
<td>29</td>
<td>24.9</td>
<td>29.8</td>
<td>14.3</td>
<td>69.0</td>
</tr>
<tr>
<td>30</td>
<td>25.0</td>
<td>29.8</td>
<td>14.5</td>
<td>69.3</td>
</tr>
<tr>
<td>38</td>
<td>25.1</td>
<td>30.2</td>
<td>14.4</td>
<td>69.7</td>
</tr>
<tr>
<td>26</td>
<td>25.2</td>
<td>30.1</td>
<td>14.4</td>
<td>69.7</td>
</tr>
<tr>
<td>22</td>
<td>--</td>
<td>30.0</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>32</td>
<td>--</td>
<td>30.4</td>
<td>--</td>
<td>--</td>
</tr>
<tr>
<td>23</td>
<td>25.0</td>
<td>30.3</td>
<td>14.7</td>
<td>70.0</td>
</tr>
<tr>
<td>28</td>
<td>25.4</td>
<td>30.0</td>
<td>14.6</td>
<td>70.0</td>
</tr>
<tr>
<td>21</td>
<td>25.7</td>
<td>30.9</td>
<td>15.1</td>
<td>71.7</td>
</tr>
<tr>
<td>27</td>
<td>26.2</td>
<td>30.9</td>
<td>15.2</td>
<td>72.3</td>
</tr>
<tr>
<td>24</td>
<td>27.0</td>
<td>31.8</td>
<td>15.6</td>
<td>74.4</td>
</tr>
<tr>
<td>31</td>
<td>27.0</td>
<td>32.0</td>
<td>15.8</td>
<td>74.8</td>
</tr>
<tr>
<td>39</td>
<td>27.0</td>
<td>32.1</td>
<td>15.7</td>
<td>74.8</td>
</tr>
<tr>
<td>41</td>
<td>27.1</td>
<td>32.3</td>
<td>15.6</td>
<td>75.0</td>
</tr>
<tr>
<td>40</td>
<td>27.0</td>
<td>32.3</td>
<td>15.8</td>
<td>75.1</td>
</tr>
</tbody>
</table>
LITERATURE CITED


This program reads in a data matrix by columns. The $i,j$'th entry is the value of variable $i$ on entity $j$. Negative and zero values are assumed to be missing. Phenetic distance between each pair of entities is calculated. The entities are clustered by maximum linkage. Output, given at the 100, 50, 20, 10 and 2 cluster stages, consists of the least distance at the preceding stage (maximum cluster diameter) and the number of the cluster to which each entity is assigned. At the 10 and 2 cluster stages the distance matrix is printed.

The following variables play important roles:

- $NE = $ number of entities to be clustered.
- $NV = $ number of variables per entity.
- $S(J) = J$'th element in the lower triangular distance matrix.
- $NEAR(I) = $ Identification number of least element in row $I$ of the distance matrix.
- $SREF(I) = $ Least element in row $I$.
- $SREFX = $ Least element in the entire array.
- $LIST(I) = I$'th cluster identification number in sequential list of current clusters.
- $MER(I) = $ cluster to which entity $I$ is assigned.
LFIND is a function for locating the distance between pairs of clusters. Distances are thought of as a 2-dimensional array but are stored as a 1-dimensional array. LFIND is the correspondence between the arrays.

Storage requirements for dimensioned variables are:

<table>
<thead>
<tr>
<th>Variable</th>
<th>Type</th>
<th>Dimension</th>
</tr>
</thead>
<tbody>
<tr>
<td>S(I)</td>
<td>Integer</td>
<td>NE(NE-1)/2</td>
</tr>
<tr>
<td>SREF(I)</td>
<td>Integer</td>
<td>NE</td>
</tr>
<tr>
<td>CH(I, J)</td>
<td>Real</td>
<td>NV \times NE</td>
</tr>
<tr>
<td>LIST(I)</td>
<td>Integer</td>
<td>NE</td>
</tr>
<tr>
<td>MER(I)</td>
<td>Integer</td>
<td>NE</td>
</tr>
</tbody>
</table>

Total core storage needed is \(\text{NE(NE-1)/2 + 4NE + 2NE \times NV}\) words, plus about 1,200 words for the program.

The user must replace cards 13 and 14 so that dimensions are appropriate to the problem. If it is anticipated that some values of \(\frac{1}{p} \sum (CH(K, I) - CH(K, J))^2\) will be greater than 63, then the user must change either card 38 or card 45 so that all values of \(S(N)\) are less than BIG. To estimate the square root, recall that it is
no larger than the maximum of \(CH(K,I) - CH(K,J)\). If the
program is not run on the CDC3300 at OSU, all READ and
WRITE statements must be changed so that input/output
device numbers are appropriate.

The first data card contains \(NE\) and \(NV\) in format
2I4. The second through \(n+1\)st data cards contain values
for the first entity, arranged in the format the user
specifies in card 21 (Format card 103). The \(n+2\)nd through
2\(n+1\)st cards contain values for the second entity, and
so on.
FUNCTION LFIND(I,J)
004 IF (I-J) 9,29,19
005 19 LFIND = ((J-1)*(J-2))/2+I
006 RETURN
007 29 LFIND = ((I-1)*(I-2))/2+J
008 RETURN
010 WRITE (5,209)
011 209 FORMAT ('ERROR...LOOKING AT MATRIX DIAGONAL IS FORBIDDEN')
012 CALL EXIT
013 END
014 PROGRAM CLUSTER
015 INTEGER S(16110), SREF(180), BIG, SREFX
016 DIMENSION CH(26,180), LIST(180), MER(180), NEAR(180)
017 C READ NUMBER OF ENTITIES AND NUMBER OF VARIABLES IN 2 FIELDS OF 4
018 READ (60,203) NE, NV
019 203 FORMAT (2I4)
020 C READ DATA MATRIX
021 DO 13 I=1,NE
022 READ (60,103) (CH(K,I), K=1,NV)
023 103 FORMAT (6F4.0,4X,2F4.0,4X,8F4.0,/,10F4.0)
024 13 CONTINUE
025 C CALCULATE DISTANCES AS ROWS OF A LOWER TRIANGULAR MATRIX AND INDEX
026 C AS A ONE DIMENSIONAL ARRAY. I = ROW INDEX, J = COLUMN INDEX, K =
027 C VARIABLE INDEX, F = NUMBER OF POSITIVE VARIABLES
028 N = 1
029 DO 53 I =2,NE
030 II = I-1
031 P = 0.
032 Y = 0.
033 DO 43 K = 1,NV
034 43 IF (CH(K,I)) 43,43,23
035 23 IF (CH(K,J)) 43,43,33
036 33 Y = Y + (CH(K,I) - CH(K,J)) **2
P = P + 1.

D = SQRT(Y/P) * 512. + .5

S(N) = D

N = N + 1

CONTINUE

CONTINUE

CONTINUE

C INITIALIZE VARIABLES AND SET CONSTANTS

NCL = NE

BIG = 32767

C INITIALIZE ARRAYS

DO 10 J = 1,NE
    LIST(J) = J
    SREF(J) = BIG
    MER(J) = J

CONTINUE

DO 10 J = 1,NE
    LIST(J) = J
    SREF(J) = BIG
    MER(J) = J

CONTINUE

C FIND MINIMUM ENTRY IN EACH ROW

L = 0

DO 30 I = 2,NCL
    I1 = I - 1
    DO 30 J = 1, I1
    L = L + 1

C IN EFFECT S(L) = S(I,J)

IF ((S(L) - SREF(I))) 31,31,30

31 NEAR(I) = J

30 CONTINUE

31 CONTINUE

C MAIN LOOP. FIND MINIMUM VALUE IN SREF ARRAY.

SREFX = BIG

DO 50 I = 2,NCL
    LISTI = LIST(I)
    IF ((SREF(LISTI) - SREFX)) 51,51,50

51 IREF = I

50 CONTINUE

51 CONTINUE

SREFX = SREF(LISTI)
071  50  CONTINUE
072  C  LREF IS THE NUMBER OF THE ROW CONTAINING THE SMALLEST ENTRY IN THE
073  C  S ARRAY. IF THERE ARE TIES, THEN LREF IS THE HIGHEST NUMBERED ROW
074  C  CONTAINING THIS SMALLEST VALUE. HENCE LREF IS LARGER THAN NEAR (LREF)
075  C  IREF IDENTIFIES THE PLACEMENT OF LREF IN THE LIST ARRAY
076                   NREF = NEAR(LREF)
077  C UPDATE MER
078       DO 20 I = 1, NE
079           IF (MER(I) - LREF) 20, 21, 20
080        21  MER(I) = NREF
081        20  CONTINUE
082       NCL = NCL - 1
083  C UPDATE LIST BY REMOVING LREF AND PUSHING DOWN LIST
084       IF (IREF - NCL) 91, 91, 90
085        91  DO 80 I = IREF, NCL
086            LIST(I) = LIST(I+1)
087        80  CONTINUE
088        90  CONTINUE
089  C UPDATE S ARRAY BY REPLACING S(I,NREF) WITH THE LARGER OF S(I,NREF)
090  C AND S(I,LREF)
091       DO 37 J = 1, NCL
092           I = LIST(J)
093          IF (I - NREF) 36, 37, 36
094  C RECALL THAT LREF HAS BEEN REMOVED FROM LIST SO I NEED NOT BE TESTED
095  C FOR EQUALITY WITH LREF
096        36  LL = LFIND(I, LREF)
097        37  LN = LFIND(I, NREF)
098          IF ((S(LL) - S(LN))) 37, 38, 37
099        38  S(LN) = S(LL)
100        37  CONTINUE
101  C UPDATE THE NEAR AND SREF ARRAYS. IF THE LEAST ELEMENT IN ROW I WAS IN
102  C EITHER COLUMN LREF OR COLUMN NREF, THEN WE MUST FIND A NEW LEAST
103  C ELEMENT. WE NEED NOT CONSIDER ROWS BEFORE NREF.
104       DO 57 J = 1, NCL
105           I = LIST(J)
IF (I-NREF) 57, 27, 57
CONTINUE
IF (J-1) 67, 87, 67
SREF(I) = BIG
J1 = J-1
DO 77 L = 1, J1
LISTL = LIST(L)
LL = LFIND(I, LISTL)
IF (NEAR(I) - SREF(I)) 76, 77, 77
CONTINUE
SREF(I) = S(LL)
CONTINUE
J = J+1
IF (J-NCL) 96, 96, 97
I = LIST(J)
IF (NEAR(I) - LREF) 64, 67, 64
IF (NEAR(I) - NREF) 66, 67, 66
GO TO 87
CONTINUE
C CHECK IF OUTPUT IS WANTED
IF (NCL .NE. 100 .AND. NCL .NE. 50 .AND. NCL .NE. 20 .AND. NCL .NE. 10 .AND. NCL .NE. 2) GO TO 40
WRITE (61, 105) NCL, SREFX
FORMAT ('ENTITY', 7X, 'CLUSTER')
WRITE (61, 115)
DO 32 J = 2, NCL
WRITE (61, 125) I, MER(I)
FORMAT ('ENTITY', 7X, 'CLUSTER')
WRITE (61, 135) NCL, NCL
CONTINUE
IF (NCL .NE. 10 .AND. NCL .NE. 2) GO TO 40
WRITE (61, 135)
FORMAT ('ENTITY', 7X, 'CLUSTER')
CONTINUE
DO 32 J = 2, NCL
I = LIST(J)
J1 = J-1
DO 42 L = 1, J1
LISTL = LIST(L)
LL = LFIND(LISTL)
WRITE (61, 145) I, LISTL, S(LL)
FORMAT (1X, I4, 8X, I4, 8X, I6)
CONTINUE
IF (NCL - 2) 40, 62, 40
CONTINUE
STOP
END