CONF-9408107--7

DIFFERENTIATION OF AROCLORS IN ENVIRONMENTAL SAMPLES USING LINEAR DISCRIMINATION

C. Y. Ma[a] and C. K. Bayne[b]
Oak Ridge National Laboratory*
P.O. Box 2008
Oak Ridge, Tennessee, 37831

[a]Chemical and Analytical Sciences Division
[b]Computing Applications Division

Presented at 1994 Joint Statistical Meetings,
Toronto, Ontario, Canada, 14-18 August 1994.

*Managed by Martin Marietta Energy Systems, Inc. under contract DE-AC05-840R21400 with the U.S. Department of Energy.

DIFFERENTIATION OF AROCLORS IN ENVIRONMENTAL SAMPLES USING LINEAR DISCRIMINATION

C. Y. Ma[a] and C. K. Bayne[b]
Oak Ridge National Laboratory, P.O. Box 2008, Oak Ridge, Tennessee, 37831
[a]Chemical and Analytical Sciences Division
[b]Computing Applications Division

Introduction

Environmental samples suspected of containing polychlorinated biphenyls (PCBs) and analyzed by Environmental Protection Agency (EPA) methods frequently contain non-PCB components, such as phthalates, polynuclear aromatic hydrocarbons, or organochlorine pesticides. The presence of these interferences can often obscure the gas chromatographic (GC) patterns and cause problems in differentiating the Aroclor types by visual inspection. Because EPA methods require the identification of Aroclor types in order to determine which Aroclor standard should be used for the quantitation calibration, and possibly, to trace the source of PCB occurrences, electron capture negative ion chemical ionization (ECNICI) mass spectrometry was used to provide additional parameters for discriminating PCB congeners from the interferences.

In this study, eight Aroclors (i.e., 1016, 1232, 1242, 1248, 1254, 1260, 1262, and 1268) were analyzed by ECNICI for a range of concentration levels. Using selective ion display, the ion abundances of the prominent peaks in the isotopic clusters of molecular ions were measured for eight PCB homologs (Cl_2-Cl_9) within their GC retention time windows for each sample. Corresponding relative ion abundances from eight Aroclor standards were used as classification training sets. These training sets were used to develop a classification algorithm using standard linear discriminant analysis to classify the PCBs present in an unknown sample(s) as a specific Aroclor. This technique employed a sequential application of two sets of linear discriminant functions to successfully identify known EPA quality control and environmental samples. However, classifications of Aroclor mixtures were less successful.

Training Set

ECNICI spectra for PCB samples were obtained with a Hewlett-Packard 5985B Gas Chromatograph/Mass Spectrometry system with dual electron impact/chemical ionization source. Replicate samples were analyzed using either two or three concentration levels on each of the eight Aroclor standard solutions. Concentration values for Aroclor 1016 were 20 and 100 ppm; for Aroclors 1232 and 1242 were 10 and 100 ppm; for Aroclor 1248 were 1 and 10 ppm; and for Aroclors 1254, 1260, 1262, and 1268 were 0.2, 1, and 10 ppm. A total of 67 samples with these different concentration levels were analyzed and used as training sets for developing a methodology to classify unknown environmental samples.

Using selective ion display, the ion abundances for the prominent peak in the molecular ion isotopic clusters were integrated for eight PCB homologs (Cl_2-Cl_9) over the PCB retention time window (10 to 50 min). The ions selected for the ion abundance measurements were m/z 221, 256, 292, 326, 360, 394, 430, and 464, corresponding to the eight PCB homologs ($C_{12}H_{10-x}Cl_x$, X=2 to 9), respectively. Relative ion abundance (RIA) was defined as the ion abundance of a homolog relative to the sum of all eight homologs. The eight RIA measurements represent one observation vector for each analyzed Aroclor sample in the classification training sets.

Response Variable Selection.

Response vectors with RIAs elements were used to classify a sample as an Aroclor type by linear discrimination functions (LDFs). First, RIAs were selected as elements in the response vector based on the performance of each individual RIA. Individual RIAs were ranked from largest influence to smallest influence by the ratio of between-to-within standard deviations. A between standard deviation for measurements of a particular RIA is an overall estimate of variability between different Aroclor training sets. A within standard deviation for measurements of a particular RIA is an overall estimate of variability for replicate measurements within each Aroclor training set. Standard deviations were estimated from variance components in a one-way analysis of variance. Ideally, RIA standard deviations *between* Aroclor training sets should be very *large*, but RIA standard deviations *within* each Aroclor training set should be very *small*. Large between-to-within

ratios of RIA standard deviations would be more useful for discriminating among Aroclor types.

Second, each RIA was added sequentially as an element to the response vector in the order of their between-to-within ratio ranking. With each addition to the response vector, the pairwise distances between all Aroclor training sets were calculated by the Mahalanobis distance function. The most difficult classification is when two Aroclor training sets have a small Mahalanobis distance. Therefore, RIAs were added to the response vector until the Mahalanobis distance between the two closest (minimum distance) Aroclor training sets of RIA measurements was near the maximum separation.

Sequential Classification

A classification algorithm using standard linear discriminant analysis was developed to classify an unknown sample(s) as a specific Aroclor type. This approach established training sets of known RIA patterns by analyzing samples of Aroclor standards. These training sets were used to estimate the coefficients for eight linear discriminant functions corresponding to the eight different Aroclor types. The classification algorithm uses the linear discriminant functions to classify unknown samples in two stages. The first stage selects the "best" RIAs from the observation vectors as elements of response vectors that are used to classify samples into Aroclors 1016, 1268, and "others." For those samples classified as "others," the second stage selects a different set of "best" RIAs to classify samples into Aroclors 1232, 1242, 1248, 1254, 1260, and 1262. This two-stage approach was developed because the Mahalanobis distances between RIA response vectors in Aroclors 1016 and 1268 training sets and the remaining Aroclor training sets were comparatively large.

A large Mahalanobis distance between training sets for Aroclors 1016 and 1268 indicates that these training sets are easily classified with the initial linear discriminant analysis. The second stage refines the linear discriminant analysis by selecting different RIAs as elements for the response vector to improve the classification of samples in the remaining Aroclors.

Classification Results

The LDFs have been used to identify independent samples of Aroclor standards that have been analyzed by ECNICI over a period of two years. Thirty analyses were classified that contain eight Aroclor standards at various concentrations (1 to 100 ppm). All Aroclor standards were identified correctly. Twenty-seven samples were identified with an identification probability of greater than 0.98; the remaining three, with probabilities greater than 0.75. In addition, coefficient of variations for LDF values are less than 10% for Aroclor standards 1254 and 1260 that have been analyzed intermittently over the last two years. This long-term stability further indicates that once the instrument is tuned according to a standard operating procedure, this technique is very reliable.

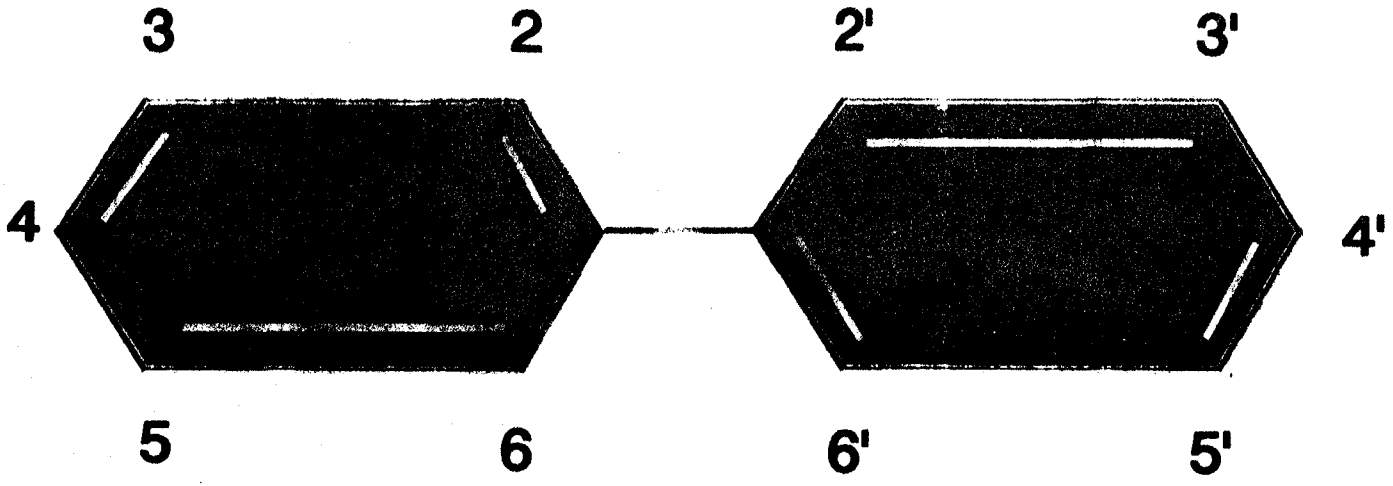
To extend the validity of the sequential classification method, the method was applied to the analysis of EPA quality control samples. Intermittently over a period of six months, six samples prepared from two EPA transformer oil samples (containing either Aroclor 1254 or Aroclor 1260) were analyzed by ECNICI and classified. The sequential classification method identified all six quality control samples correctly with an identification probability of one.

Twenty-two environmental samples in various matrices, including soil, water, water sludge, clam tissue, fish, oil paint, waste oil and heat transferline oil samples were analyzed. The presence of PCBs in those samples was first detected by GC/ECD. However, because their GC/ECD profiles were obscured by the presence of non-PCB components, classification of the Aroclor types in some of the samples by merely visual inspection (as required by EPA Methods) was very difficult, if not impossible. Subsequently, all these environmental samples were analyzed by ECNICI, and the PCB components present in the samples were identified by the LDFs. Except for four samples, the remaining eighteen were identified with an identification probability of greater than 0.97. Presumably, the four unclassified samples contain either a mixture of Aroclors, or Aroclors altered by weathering, by metabolic and/or chemical degradation.

Differentiation of Aroclors in Enviromental Samples Using Linear Discrimination

C. Y. Ma and C. K. Bayne
Oak Ridge National Laboratory

209 Different PCBs

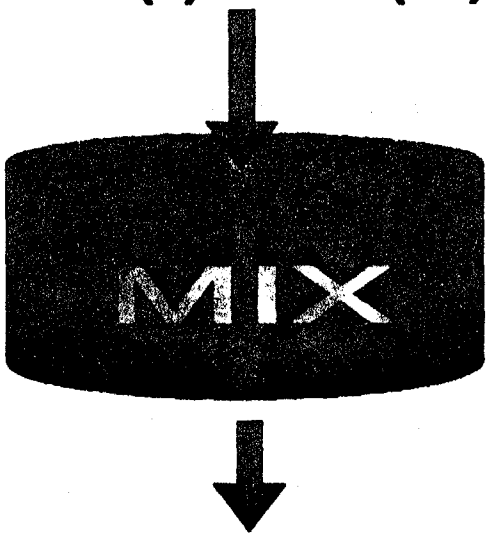


PCB(1) = 3
PCB(2) = 12
PCB(3) = 24
PCB(4) = 42
PCB(5) = 46

PCB(6) = 42
PCB(7) = 24
PCB(8) = 12
PCB(9) = 3
PCB(10) = 1

Aroclors are PCB Mixtures by Weight

PCB(1) - PCB(10)

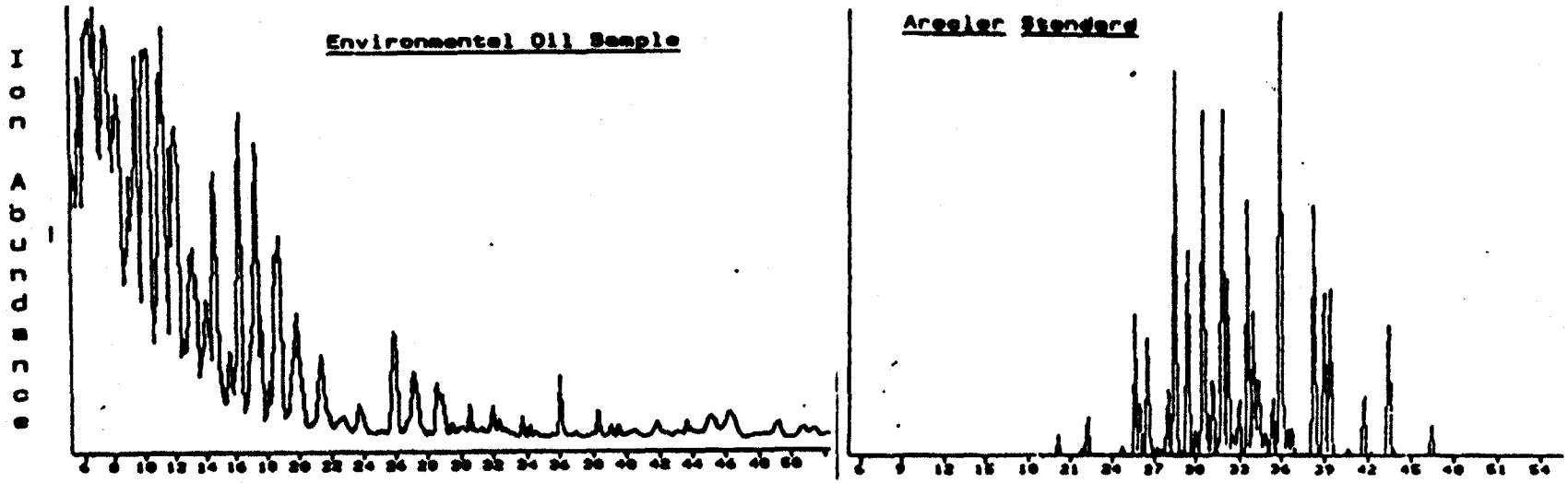


Aroclors

1016	1242	1260
1221*	1248	1262
1232	1254	1268

Interference from Non-PCB Components

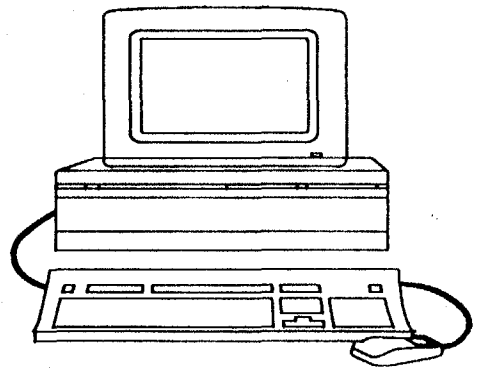
PROBLEM: Sample Components Obsure GC/ECD Pattern



SOLUTION: NICI Provides additional Information

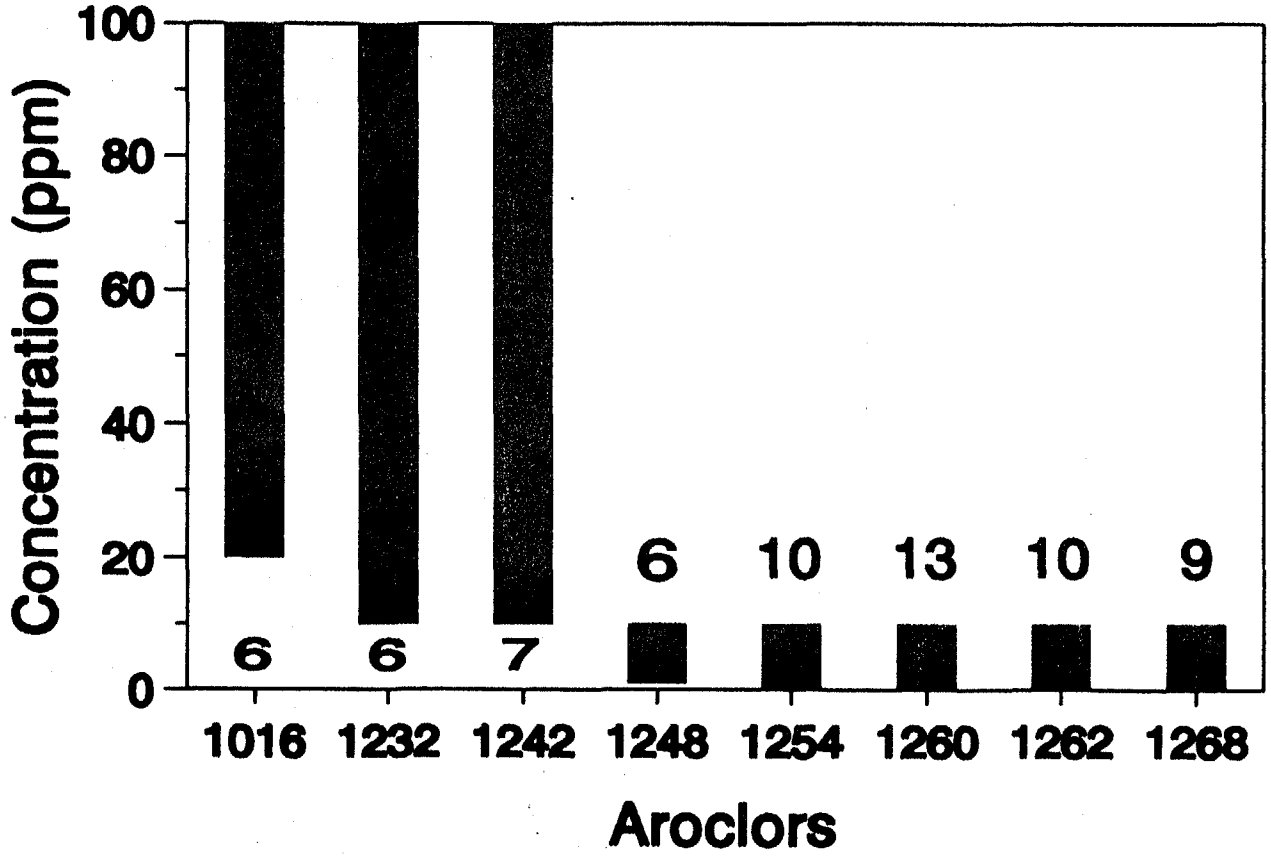
Mass Spectrometric Analysis Measures Aroclor Masses

Aroclor 1248

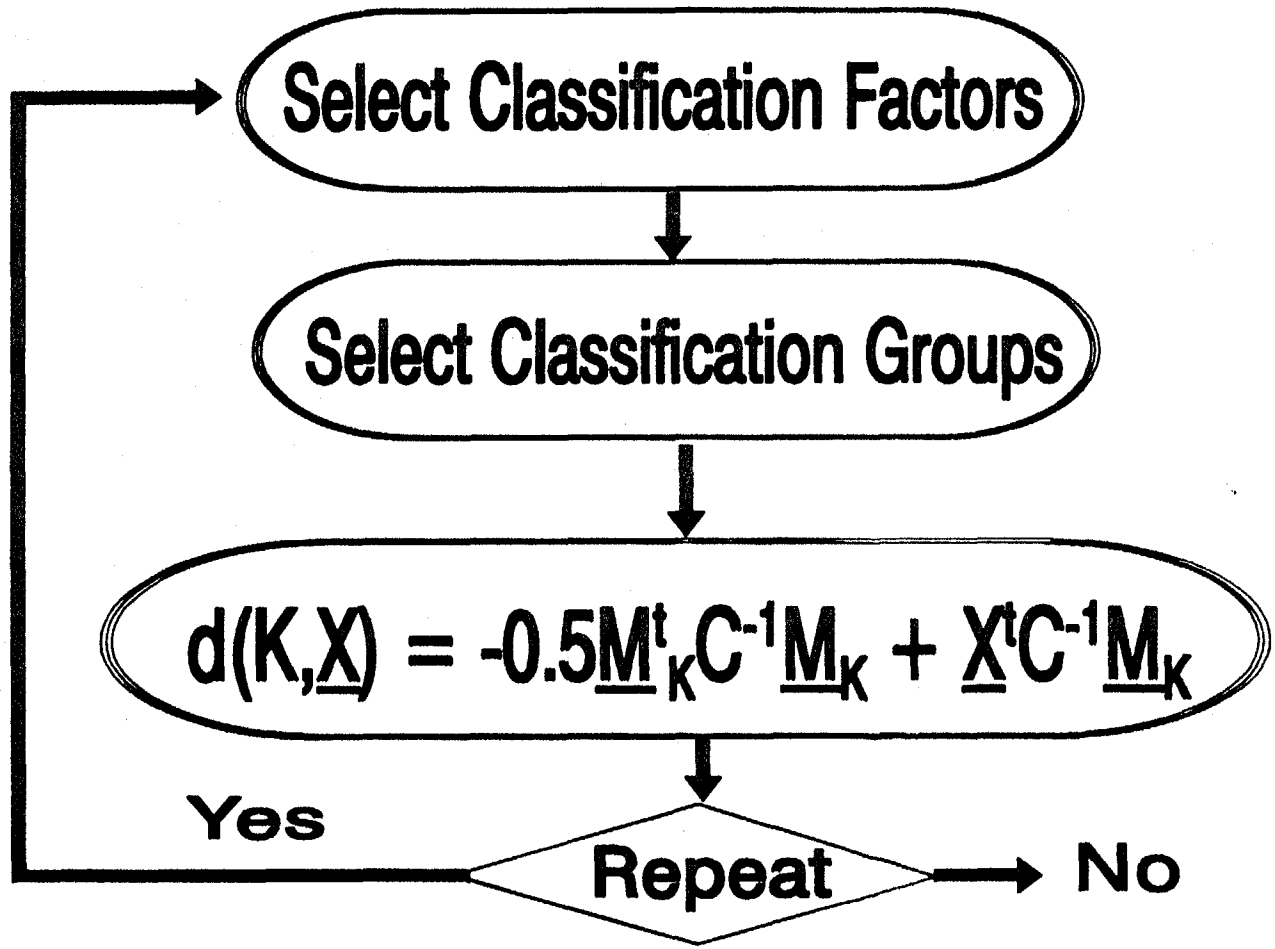


Mass	% Relative Abundance
221	3%
256	4%
292	26%
326	34%
360	11%
394	14%
430	5%
463	3%

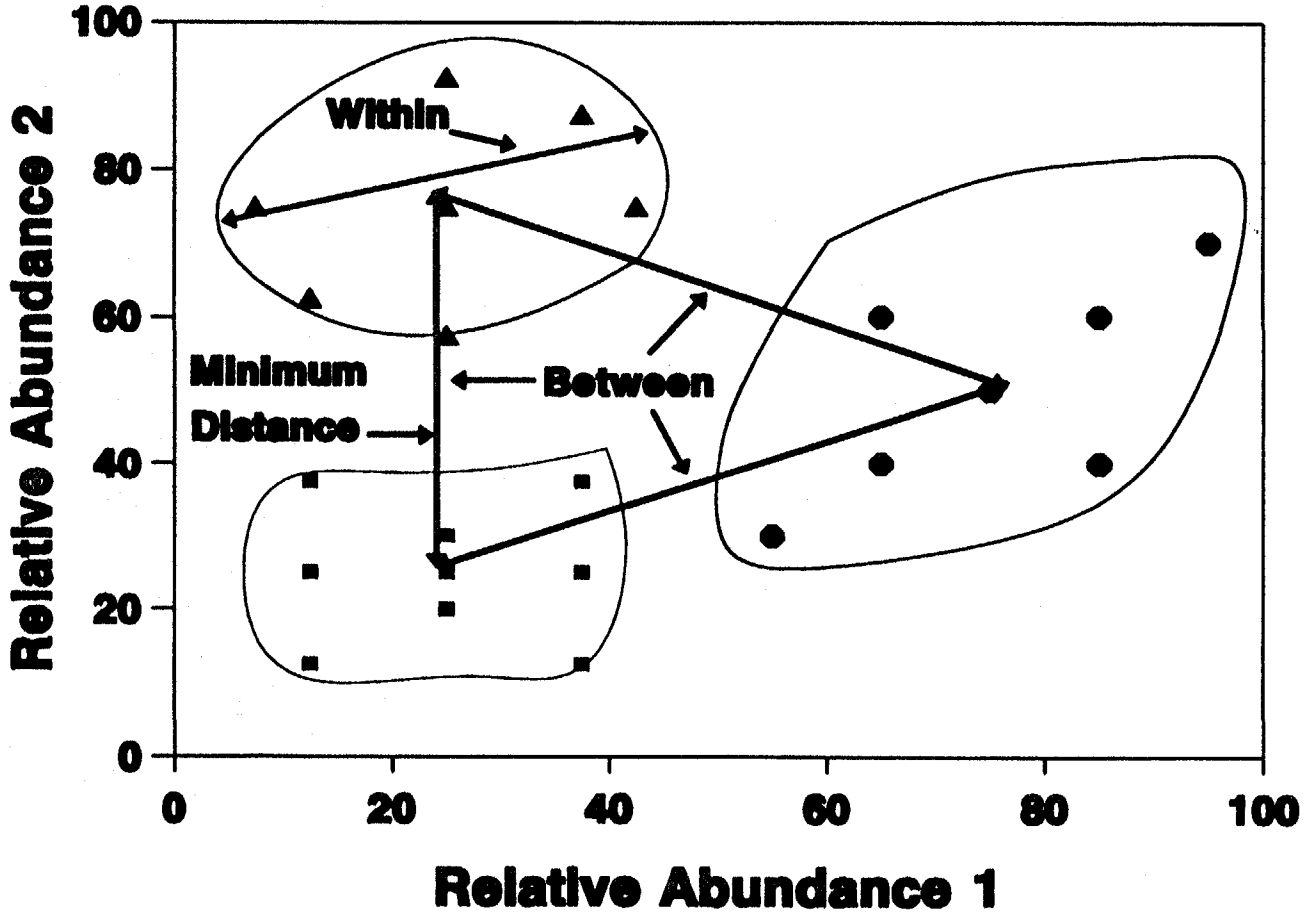
67 Aroclor Standards Formed the Training Set



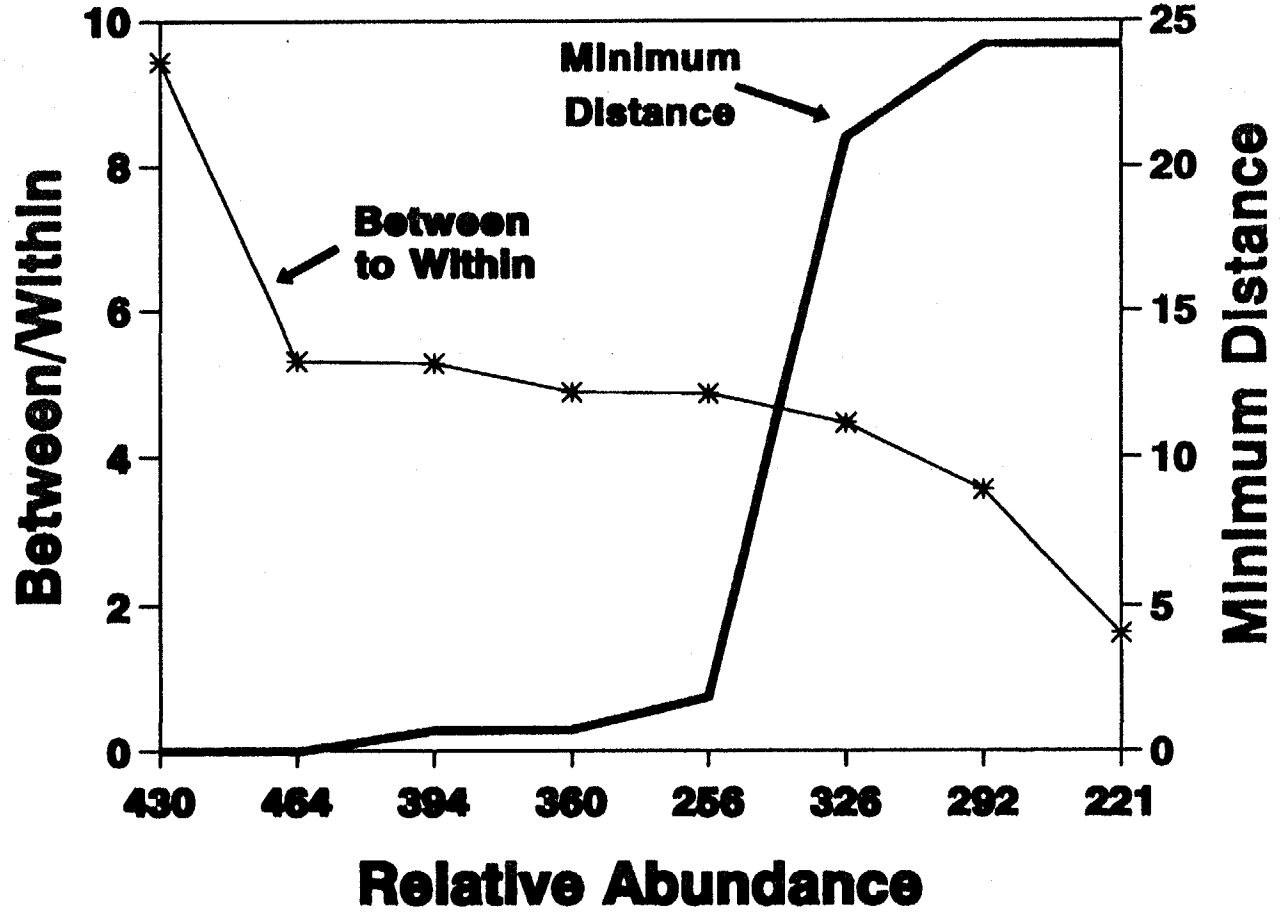
Sequential Classification Method



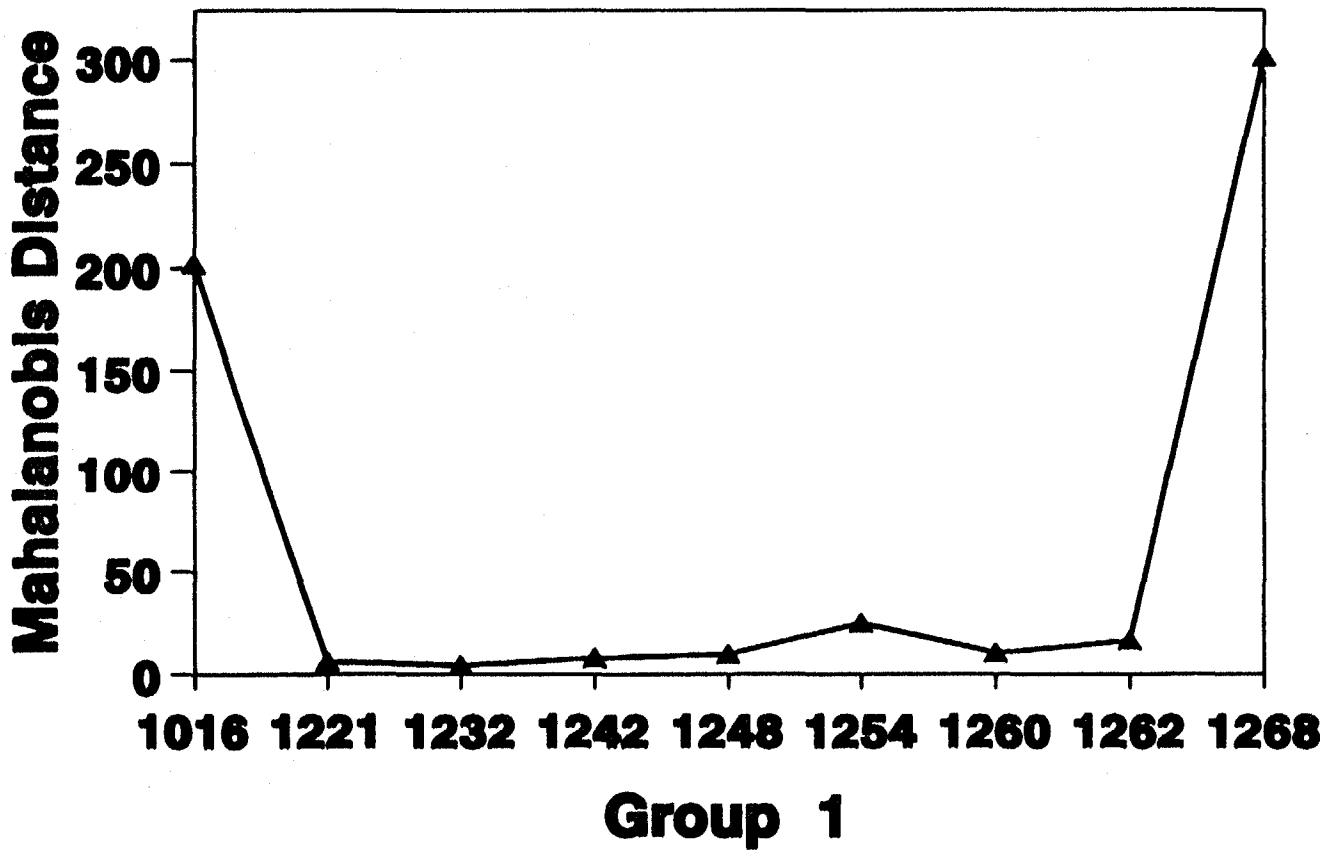
Between and Within Distances



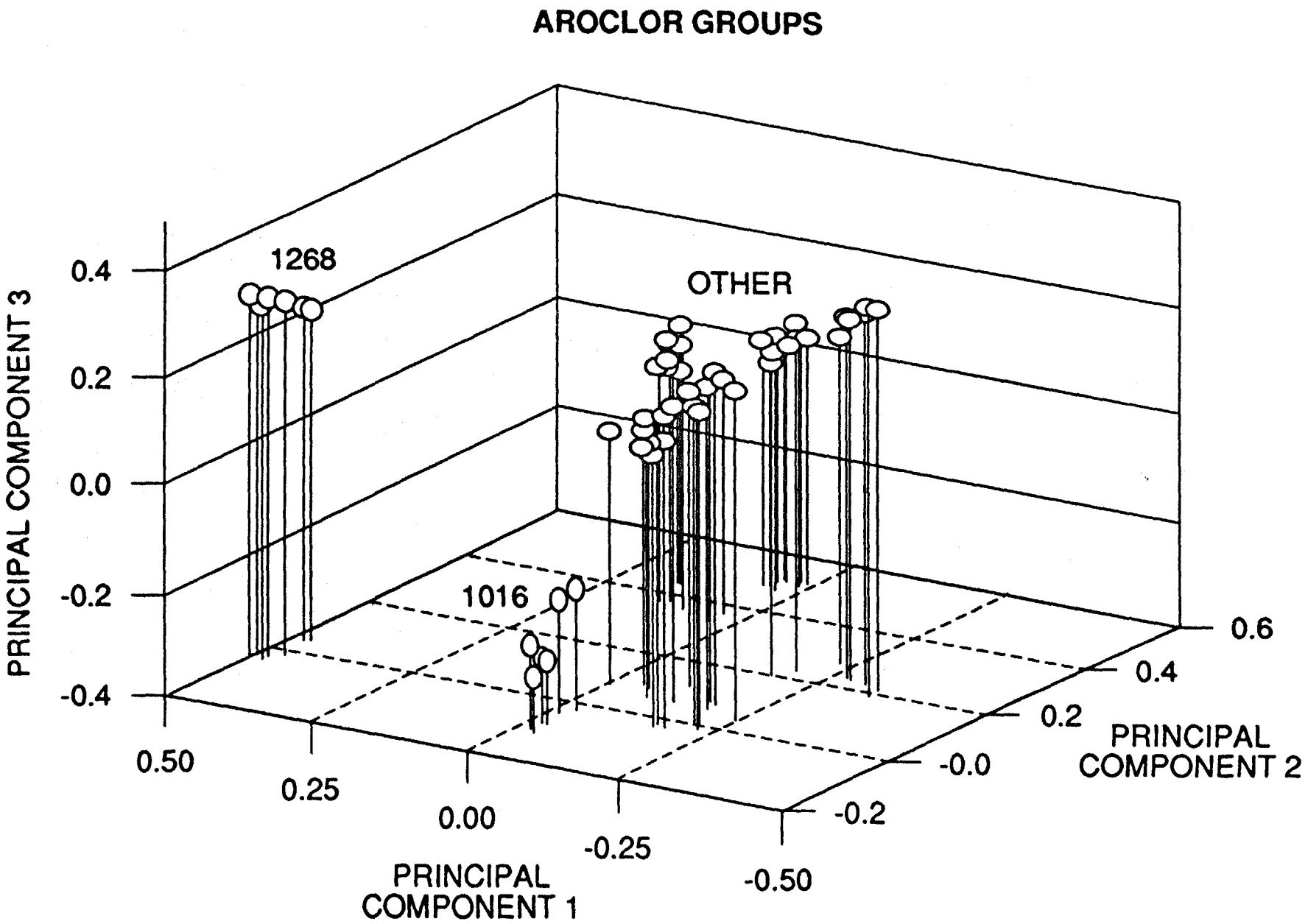
Select Classification Factors



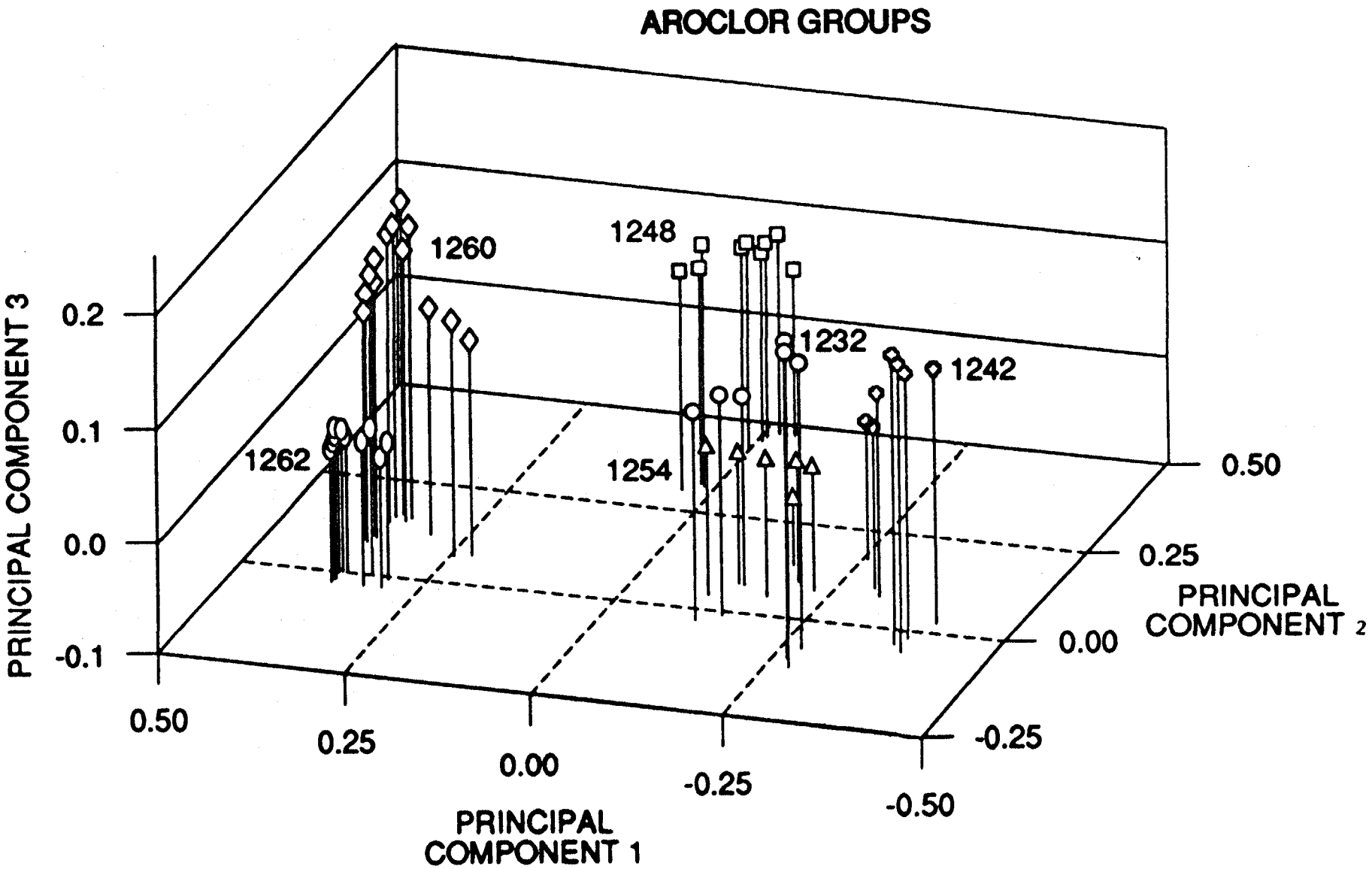
Select Classification Groups
Group 1 to Remaining Aroclors



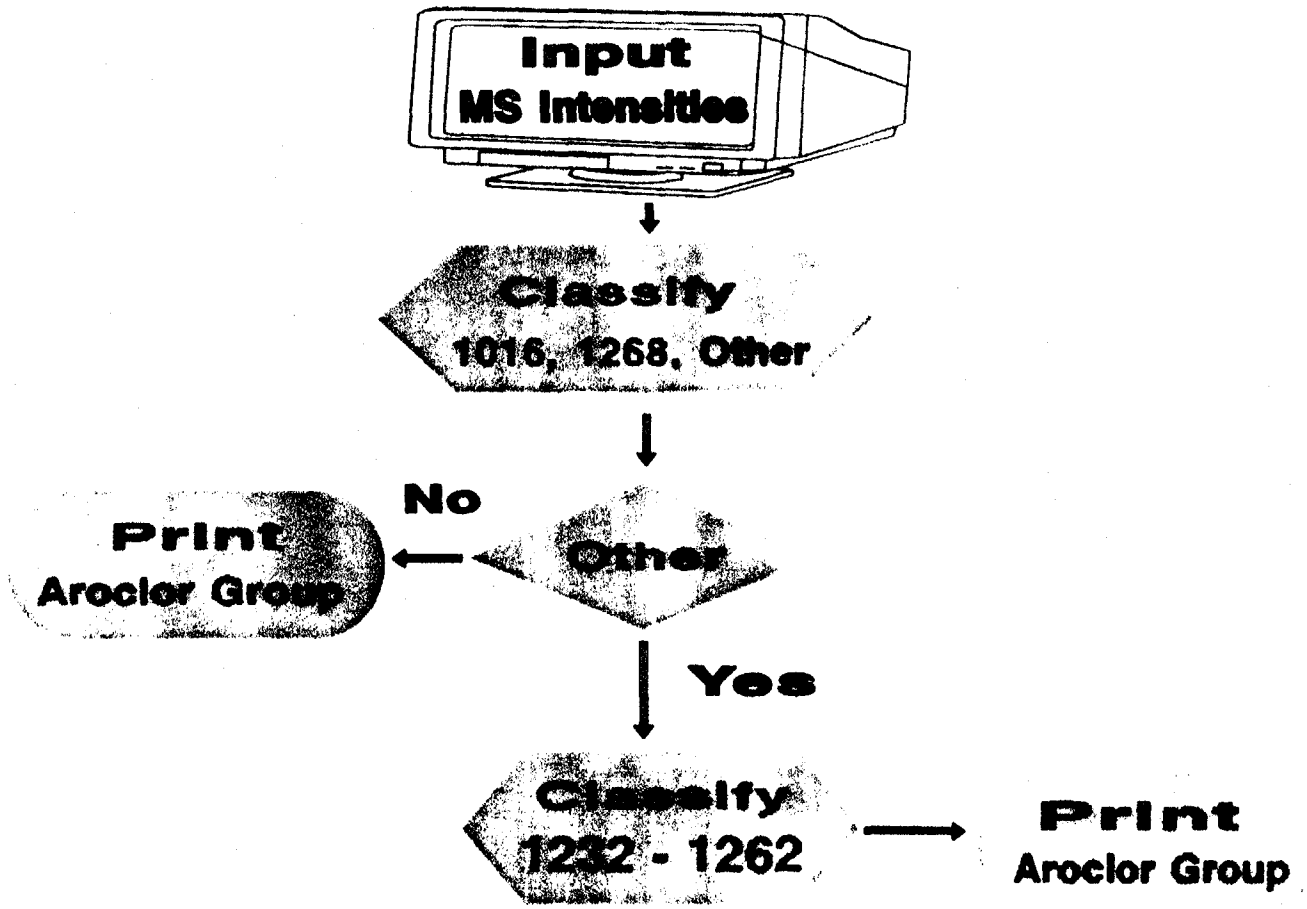
ORNL-DWG 92M-7982



ORNL-DWG 92M-7983



Sequential Classification Method



Performance of SCM

- SCM Correctly Classified

  6 EPA Transformer Oil Samples

  30 Samples of 8 Different Aroclor Standards

- Used as a QC Check for the Past 2 Years

- Unable to Classify Mixtures