Arch. Environ. Contam. Toxicol. 20, 183–199 (1991)

Archives of Environmental Contamination and Toxicology

Chemometric Comparison of Polychlorinated Biphenyl Residues and Toxicologically Active Polychlorinated Biphenyl Congeners in the Eggs of Forster's Terns (*Sterna fosteri*)

Ted R. Schwartz and David L. Stalling[1]

National Fisheries Contaminant Research Center, U.S. Fish and Wildlife Service, 4200 New Haven Road, Columbia, Missouri 65201, USA

Abstract. The separation and characterization of complex mixtures of polychlorinated biphenyls (PCBs) is approached from the perspective of a problem in chemometrics. A technique for quantitative determination of PCB congeners is described as well as an enrichment technique designed to isolate only those congener residues which induce mixed aryl hydrocarbon hydroxylase enzyme activity. A congener-specific procedure is utilized for the determination of PCBs in which *n*-alkyl trichloroacetates are used as retention index marker compounds. Retention indices are reproducible in the range of ± 0.05 to ± 0.7 depending on the specific congener. A laboratory data base system developed to aid in the editing and quantitation of data generated from capillary gas chromatography was employed to quantitate chromatographic data. Data base management was provided by computer programs written in VAX-DSM (Digital Standard MUMPS) for the VAX-DEC (Digital Equipment Corp.) family of computers.

In the chemometric evaluation of these complex chromatographic profiles, data are viewed from a single analysis as a point in multi-dimensional space. Principal Components Analysis was used to obtain a representation of the data in a lower dimensional space. Two- and three-dimensional projections based on sample scores from the principal components models were used to visualize the behavior of Aroclor® mixtures. These models can be used to determine if new sample profiles may be represented by Aroclor profiles. Concentrations of individual congeners of a given chlorine substitution may be summed to form homologue concentration. However, the use of homologue concentrations in classification studies with environmental samples can lead to erroneous conclusions about sample similarity. Chemometric applications are discussed for evaluation of Aroclor mixture analysis and compositional description of environmental residues of PCBs in eggs of Forster's terns (*Sterna fosteri*) collected from colonies near Lake Poygan and Green Bay, Wisconsin. The application of chemometrics is extended to the comparison of: a) Aroclors and PCB-containing environmental samples; to b) fractions of Aroclors and of environmental samples that have been enriched in congeners which induce mixed aryl hydrocarbon hydroxylase enzyme activity.

Polychlorinated biphenyls (PCBs) constitute a complex heterogeneous group having 209 possible congeners distributed among Cl_{1-10} homologues. PCBs have been produced by several industries worldwide in the form of technical formulations (Hutzinger *et al.* 1974). Most PCBs produced in the United States originated as one of several products designated as Aroclors®, and were manufactured by the Monsanto Chemical Co (Brinkman and de Kok 1980). The major materials produced were Aroclors 1242, 1248, 1254, and 1260. The last two digits designate the percentage of chlorine by weight in the Aroclor. Each Aroclor is characterized by a different distribution of homologues and congeners having a chromatographic profile of about 100 to 150 constituents (Ballschmiter and Zell 1980; Albro *et al.* 1981; Bush *et al.* 1982).

It is difficult to identify and quantify the individual congeners in technical formulations of PCB containing Aroclors and environmental PCB residues derived from these materials. In spite of the concern about contamination with PCBs since their discovery as environmental pollutants by Jensen (1966), much remains to be determined about their ultimate effects and fates in the environment. This lack of knowledge is due in part to the complexity of the chromatographic profile and the associated problems that must be overcome in data reduction and interpretation.

The interpretation of PCB residue data is challenging from several perspectives: (1) the data obtained from a single analysis are numerous (*e.g.*, 100–150 PCB congeners are commonly observed in a single environmental sample), (2) little understanding of the environmental distribution of PCBs can be obtained from a single analysis, (3) source profiles of PCB input into the environment are poorly characterized, (4) PCB congeners in the original polluting material often merge with congeners from other sources, and (5) the contaminant mixture may be altered by metabolism, and be-

[1] Current address, Analytical Bio-Chemistry Laboratories, Inc., P.O. Box 1097, Columbia, MO 65205.

come partitioned into multiple environmental compartments that may be further changed by weathering or degradation. A thorough understanding of these processes and correlation of residue profiles with specific toxic responses requires congener specific methods of analysis and multivariate statistical tools.

Until recently, PCBs have been quantitated by comparing selected peak areas observed in samples with those in one of several Aroclor mixtures (Webb and McCall 1973; USEPA 1979). Packed column gas chromatography has usually been used in these analyses, even though this technique provides poor resolution of individual isomers and congener groups (Duinker *et al.* 1980). The problems associated with characterizing metabolically altered or weathered PCBs is a formidable task that requires an enhanced analytical approach.

The development of a congener specific method that can provide detailed information of the many congeners present in the environment and biota has been a goal of many environmental chemists (Ballschmiter and Zel 1980; Brinkman and de Kok 1980; Rappe and Buser 1980; Schwartz *et al.* 1984). Difficulty in characterizing these residues is compounded by the massive amounts of data generated by high resolution capillary gas chromatography (GC) (Stalling *et al.* 1985a) and by the lack of qualitative and quantitative analytical standards (Mullin *et al.* 1984).

Progress has been made, however, in analyzing PCBs in many substrates. Erickson (1986) has reviewed many of these methods and many isolation and sample cleanup references are summarized in his review. As stated by Erickson (1986), the analysis of PCBs is frequently improved by GC separation using capillary columns and quantitation of the congeners by electron capture or mass spectrometric detection. Calibration and quantitation can be based on the use of one or more of five approaches:

(1) molar response factors determined with a flame ionization detector;
(2) internal standards with relative response factors;
(3) external standards of individual congeners;
(4) Aroclors or their mixtures;
(5) use of an isotopically labeled member of each homologue and extrapolation of its response to congeners of each homologue (as is often done in GC/MS analyses).

Special concern is being focused on PCB congeners which have high binding affinity to hepatic cytosolic receptor protein (Ah receptor) and high induction potency for 3-methylcholanthrene (MC) type hepatic microsomal enzymes. Several of these congeners are isosteric with 2,3,7,8-tetrachlorodibenzo-*p*-dioxin (Polland and Knutson 1982; Bandiera *et al.* 1984). Within the PCB congener group there are a limited number of compounds congeners which elicit toxic responses such as porphyria, teratogenesis, endocrine and reproductive dysfunction, and lymphoid involution (McConnell 1980). Common to this group of congeners is chlorine substitution in both para positions, in two or more of the four meta positions, and in less than three of the four ortho positions. If this concept of the toxicological significance of these relatively few congeners is accurate, then characterization of PCB mixtures or environmental residues should be centered on these few compounds. However, current analytical methods do not permit the determination or enrichment of the PCB congeners which induce aryl hydrocarbon hydroxylase (AHH) or ethoxyresorufin O-deethylase (EROD) enzymatic activity. Because these biologically active PCB congeners are present at low concentration, complicated by the much greater concentration of other congeners which elute near or coelute during gas chromatographic analyses, determination of the AHH-active congeners requires enrichment and separation. We have developed such a technique based upon activated carbon dispersed on glass fibers (Smith 1981), which separates these most enzymatically active PCB congeners.

After data acquisition and quantitation, a most important step remains. The data must be examined for quality control and information content. The problem of data reduction and interpretation is approached from a chemometric viewpoint (Stalling *et al.* 1985b). Data from the analysis of many samples form a data cluster that may have structure related to such factors as exposure or distance from discharge. In the chemometric evaluation of complex profiles, data from a single analysis are viewed as a point in multi-dimensional space. Principal Components Analysis (PAC), a multivariate data analysis technique, can be applied to explore the data structure in this higher dimensional space (Wold and Sjöström 1977). The power of principal components modeling of multivariate data is in the graphical examination of data.

Principal component models project data from multidimensional space onto lower dimensional space (generally two or three dimensions) in a way that preserves the maximum amount of variance and relations among samples and variables (Wold *et al.* 1984a, 1984b). Principal components modeling gives good results even when the number of variables exceeds the number of samples. This technique is especially useful in visualizing sample similarities from PCB congener-specific analysis. In these analyses there are often more than 100 PCB constituents in a sample. SIMCA (Soft Independent Modeling of Class Analogy) pattern recognition technique developed by Albano and Wold and their coworkers (Alano *et al.* 1978; Wold, 1982; Wold *et al.* 1984b) is based on derivation of disjoint principal component models. These models can be used for graphical representation and classification of new samples. SIMCA has been applied to a variety of chemical problems (Wold 1984b). Using this method, Dunn *et al.* (1984) and Stalling *et al.* (1985a, 1985b) examined similarities in the composition of PCB mixtures and Aroclors. They demonstrated that three-term principal components models of Aroclor and Aroclor mixtures formed a tetrahedron-like volume in concentration space, in which mixtures of any two Aroclors formed the edge boundaries and mixtures of any three Aroclors formed the surface planes of the tetrahedron. Mixtures of four Aroclors were contained in the interior space of the tetrahedron.

Dunn *et al.* (1984) demonstrated that partial least squares (PLS) in latent variables (Sjöström *et al.* 1983; Wold 1984a) is a suitable method for determining the composition of Aroclor mixtures in samples composed of, or derived from, mixtures of Aroclors. However, before PLS is used for this purpose, classification studies are recommended to establish that residues in the samples analyzed are, or can be, accurately represented by Aroclors (Stalling *et al.* 1985b; Schwartz *et al.* 1986).

Luotamo *et al.* (1988) evaluated quantitation of PCB

levels and characterization of congener specific EC response profiles in serum specimens using both direct quantitation or generalized pattern recognition methods (partial least squares projection to latent structures) using three standard PCB reference materials. They found different serum profiles existed among several exposure groups (accidental and chronic occupational exposure and only dietary exposure). Correlation between quantitative analysis results and PLS quantitation was excellent but differences in absolute concentrations existed (slopes and intecepts of regression lines of the different quantitation methods). These differences in quantitation were attributed to the existence of different profiles in the exposed groups. Their work emphasized the need for selecting a homogeneous sample group for quantitation and they found no systematic differences between quantitation based on congener specific quantitation or PLS prediction.

Because it is difficult to statistically evaluate large sets of sample data composed of many individual PCB concentration measurements of the data are not readily available in machine-readable form. Onuska *et al.* (1985) used SIMCA to examine PCB residue profiles composed of Cl_{1-10} congener sums derived from the analysis of Aroclors and their mixtures. Onuska focused on characterizing Aroclor mixtures by using Cl_{1-10} congener profiles, and obtained principal components score plots that were similar to those obtained by modeling the concentrations of 69 congeners by Stalling *et al.* (1985b).

Schwartz *et al.* (1986) used SIMCA to assess the similarity of residue profiles in fish and turtles to Aroclors based on class models derived from congener specific PCB residue profiles (105 GC peaks). They showed that the environmental PCB residues could not adequately be described by an Aroclor or Aroclor mixture and that it would be inappropriate to report the PCB residue profiles as such. Stalling *et al.* (1987) examined the data reported by Schwartz *et al.* (1986) and examined the consequences of modeling Cl_{1-10} congener sums and the decrease in information content resulting from the congener summation into individual subgroups. The use of Cl_{1-10} homologue concentrations may be preferable when sample residue profiles are dissimilar to those of Aroclors or technical formulations. However, considerable information is lost in the calculation of homologue sums, and homologue profiles should not be used for pattern recognition studies to determine similarity to Aroclors (Stalling *et al.* 1987). They also explored the use of three-dimensional (3-D) graphics in principal components modeling.

Although classification studies can show whether sample profiles can be represented by Aroclor profiles, further complications are encountered in describing residue profiles when more than one Aroclor mixture is encountered in an ecosystem. It is important to consider not only the total PCB concentration in a sample but to characterize the distribution of individual PCB congeners in that sample. This characterization becomes especially important in studies designed to determine biological importance of specific congeners or when long-term monitoring of PCBs is considered.

Only a few examples of detailed congener-specific analyses of Aroclor and environmental residues accompanied by discussion of capillary column gas chromatography data and pattern recognition are found in the literature. Therefore, we have included a discussion of GC data from the analyses of Aroclors and their mixtures, and environmental residues of PCB present in eggs of Forster's terns.

Analysis of Environmental PCB Residues

The analysis and intrepretation of PCB residues in Forster's terns from two locations in Wisconsin provide a good example of using a chemometrics approach to data analysis by pattern recognition. Pollution of Green Bay with PCBs is associated with the presence of PCBs in recycled paper and industrial and domestics wastes. A decline in reproductive success had been observed in colonies of Forster's terns and common terns in colonies in Green Bay, Wisconsin, but not an inland colony at Lake Poygan, Wisconsin. Cormorants in some colonies from Green Bay also showed a crossed-bill syndrome and some herons had been found dead or moribund (Kubiak *et al.* 1989). Congener-specific PCB residue data from Forster's tern eggs were examined with SIMCA, which makes no a priori assumptions of sample similarity to Aroclors, to obtain a geometric representation of residue profiles from each of the two colonies and in technical Aroclors. Extracts were prepared and the residue profiles examined of only those PCB congeners known to elicit AHH-activity (Goldstein 1980; Safe *et al.* 1982, Safe *et al.* 1985, Smith *et al.* 1986, Schwartz and Smith 1987).

Sample Analysis

Extraction and Isolation of PCBs

Samples of Forster's tern eggs were sent to the National Fisheries Contaminant Research Center for congener specific PCB analysis. Details of egg collection and geographical location were given by Kubiak *et al.* (1989). The egg samples were homogenized and extracted by column chromatography and gel permeation enrichment techniques (Stalling *et al.* 1972) and PCB residues were isolated by adsorption column chromatography on silica gel followed by sulfuric acid-silica gel (Schwartz and Lehmann 1982). Separation of AHH-active congeners from the total mixture of congeners was accomplished by application of an aliquant from silica gel procedure onto the carbon/glass fibers column prepared as described previously (Smith 1981). The resulting 21 cm × 1 cm (id) glass column contained 300 mg of sieved (2–10 μm) AMOCO PX-21 carbon dispersed on 4.5 g of Whatman GF/D filter material. The column was washed sequentially with 100 ml of each of 10% and 30% dichloromethane in hexane and then in the reverse flow direction with 50 ml toluene. The bulk of the PCB congeners were collected in the first fraction. Approximately 30 congeners, about 80% of which are di-ortho chlorine substituted, were recovered in the second fraction (30%) and only non- and mono-ortho chlorine substituted congeners were recovered in the toluene fraction. The congener composition of the 30% and toluene fractions is shown in Table 2. Determinations of 18 AHH-active congeners are made by capillary GC/EC on the 30% and toluene fractions from the carbon/glass-fiber procedure.

Table 1. GC peak number, retention indices, IUPAC number and structure for 113 PCBs separated on DB-5 column from Aroclor® 1242:1248:1254:1260 mixture (Figure 1, A–D)

Peak #	Retention index #	IUPAC	Number of Cl	Number of o,o'Cl	Chemical structure
1	720.10	001	1	1	2-chlorobiphenyl
2	842.26	004	2	2	2,2'-dichlorobiphenyl
		010	2	2	2,6-dichlorobiphenyl
3	899.77	007	2	1	2,4-dichlorobiphenyl
4	923.35	006	2	1	2,3'-dichlorobiphenyl
5	935.81	005	2	1	2,3-dichlorobiphenyl
		008	2	1	2,4'-dichlorobiphenyl
6	979.09	019	3	3	2,2',6-dichlorobiphenyl
7	997.48	030	(Internal Standard)		2,4,6-trichlorobiphenyl
8	1019.69	012	2	0	3,4-dichlorobiphenyl
9	1022.50	013	2	0	3,4'-dichlorobiphenyl
10	1031.22	018	3	2	2,2',5-trichlorobiphenyl
11	1034.01	017	3	2	2,2'-4-trichlorobiphenyl
12	1051.72	024	3	2	2,3,6-trichlorobiphenyl
13	1053.23	027	3	2	2,3',6-trichlorobiphenyl
14	1069.64	016	3	2	2,2',3-trichlorobiphenyl
15	1071.25	032	3	2	2,4',6-trichlorobiphenyl
16	1098.27	029	3	1	2,4,5-trichlorobiphenyl
17	1112.05	026	3	1	2,3',5-trichlorobiphenyl
18	1116.06	025	3	1	2,3',4-trichlorobiphenyl
19	1129.00	028	3	1	2,4,4'-trichlorobiphenyl
		031	3	1	2,4',5-trichlorobiphenyl
20	1145.98	021	3	1	2,3,4-trichlorobiphenyl
21	1150.11	033	3	1	2',3,4-trichlorobiphenyl
22	1152.92	053	4	3	2,2',5,6'-tetrachlorobiphenyl
23	1162.12	051	4	3	2,2',4,6'-tetrachlorobiphenyl
24	1165.58	022	3	1	2,3,4'-tetrachlorobiphenyl
25	1175.70	045	4	3	2,2',3,6-tetrachlorobiphenyl
26	1192.28	046	4	3	2,2',3,6'-tetrachlorobiphenyl
27	1206.58	052	4	2	2,2',5,5'-tetrachlorobiphenyl
28	1211.14	043	4	2	2,2',3,5-tetrachlorobiphenyl
29	1214.63	049	4	2	2,2',4,5'-tetrachlorobiphenyl
30	1219.05	047	4	2	2,2',4,4'-tetrachlorobiphenyl
31	1220.23	048	4	2	2,2',4,5-tetrachlorobiphenyl
32	1246.84	044	4	2	2,2',3,5'-tetrachlorobiphenyl
33	1252.61	042	4	2	2,2',3,4'-tetrachlorobiphenyl
34	1269.17	041	4	2	2,2',3,4-tetrachlorobiphenyl
		071	4	2	2,3',4',6-tetrachlorobiphenyl
35	1271.16	064	4	2	2,3,4',6-tetrachlorobiphenyl
36	1286.50	040	4	2	2,2',3,3'-tetrachlorobiphenyl
37	1300.45	—	4	—	—
38	1312.23	063	4	1	2,3,4',5-tetrachlorobiphenyl
39	1319.49	074	4	1	2,4,4',5-tetrachlorobiphenyl
40	1328.47	070	4	1	2,3',4',5-tetrachlorobiphenyl
		076	4	1	2',3,4,5-tetrachlorobiphenyl
41	1331.42	066	4	1	2,3',4,4'-tetrachlorobiphenyl
42	1334.24	095	5	3	2,2',3,5',6-pentachlorobiphenyl
43	1336.61	088	5	3	2,2',3,4,6-pentachlorobiphenyl
44	1348.55	091	5	3	2,2',3,4',6-pentachlorobiphenyl
45	1367.21	056	4	1	2,3,3',4'-tetrachlorobiphenyl
		060	4	1	2,3,4,4'-tetrachlorobiphenyl
46	1373.80	089	5	3	2,2',3,4,6'-pentachlorobiphenyl
47	1376.96	084	5	3	2,2',3,3',6-pentachlorobiphenyl
48	1381.03	—	5	—	—
49	1384.97	101	5	2	2,2',4,5,5'-pentachlorobiphenyl
50	1394.39	099	5	2	2,2',4,4',5-pentachlorobiphenyl
51	1406.18	119	5	2	2,3',4,4',6-pentachlorobiphenyl
52	1415.61	083	5	2	2,2',3,3',5-pentachlorobiphenyl
53	1425.55	097	5	2	2,2',3',4,5-pentachlorobiphenyl
54	1433.25	081	4	0	3,4,4',5-tetrachlorobiphenyl
55	1436.16	087	5	2	2,2',3,4,5'-pentachlorobiphenyl
56	1443.71	085	5	2	2,2',3,4,4'-pentachlorobiphenyl
57	1448.43	136	6	4	2,2',3,3',6,6'-hexachlorobiphenyl
58	1454.09	077	4	0	3,3',4,4'-tetrachlorobiphenyl

Table 1. (cont'd)

Peak #	Retention index #	IUPAC	Number of Cl	Number of o,o'Cl	Chemical structure
59	1455.76	110	5	2	2,3,3',4',6-pentachlorobiphenyl
60	1476.11	082	5	2	2,2',3,3',4-pentachlorobiphenyl
61	1480.75	151	6	3	2,2',3,5,5',6-hexachlorobiphenyl
62	1489.24	135	6	3	2,2',3,3',5,6'-hexachlorobiphenyl
		144	6	3	2,2',3,4,5',6-hexachlorobiphenyl
63	1492.20	124	5	1	2',3,4,5,5'-pentachlorobiphenyl
64	1495.52	147	6	3	2,2',3,4',5,6-hexachlorobiphenyl
65	1499.11	107	5	1	2,3,3',4',5-pentachlorobiphenyl
66	1503.08	123	5	1	2',3,4,4',5-pentachlorobiphenyl
67	1504.44	149	6	3	2,2',3,4',5',6-hexachlorobiphenyl
68	1508.14	118	5	1	2,3',4,4',5-pentachlorobiphenyl
69	1522.65	134	6	3	2,2',3,3',5,6-hexachlorobiphenyl
70	1527.64	114	5	1	2,3,4,4',5-pentachlorobiphenyl
71	1530.50	131	6	3	2,2',3,3',4,6-hexachlorobiphenyl
72	1533.36	122	5	1	2',3,3',4,5-pentachlorobiphenyl
73	1542.62	146	6	2	2,2',3,4',5,5'-hexachlorobiphenyl
74	1554.00	132	6	3	2,2',3,3',4,6'-hexachlorobiphenyl
75	1555.62	153	6	2	2,2',4,4',5,5'-hexachlorobiphenyl
76	1558.03	105	5	1	2,3,3',4,4'-pentachlorobiphenyl
77	1577.75	141	6	2	2,2',3,4,5,5'-hexachlorobiphenyl
78	1590.10	137	6	2	2,2',3,4,4',5-hexachlorobiphenyl
79	1592.24	176	7	4	2,2',3,3',4,6,6'-heptachlorobiphenyl
80	1595.78	130	6	2	2,2',3,3',4,5'-hexachlorobiphenyl
81	1605.88	138	6	2	2,2',3,4,4',5'-hexachlorobiphenyl
82	1610.07	158	6	2	2,3,3',4,4',6-hexachlorobiphenyl
83	1619.85	129	6	2	2,2',3,3',4,5-hexachlorobiphenyl
84	1621.54	126	5	0	3,3',4,4',5-pentachlorobiphenyl
85	1623.31	178	7	3	2,2',3,3',5,5',6-heptachlorobiphenyl
86	1627.94	166	6	2	2,3,4,4',5,6-hexachlorobiphenyl
87	1633.10	175	7	3	2,2',3,3',4,5',6-heptachlorobiphenyl
88	1639.49	182	7	3	2,2',3,4,4',5,6'-heptachlorobiphenyl
		187	7	3	2,2',3,4',5,5',6-heptachlorobiphenyl
89	1648.11	183	7	3	2,2',3,4,4',5',6-heptachlorobiphenyl
90	1656.37	128	6	2	2,2',3,3',4,4'-hexachlorobiphenyl
91	1661.56	167	6	1	2,3',4,4',5,5'-hexachlorobiphenyl
92	1666.95	185	7	3	2,2',3,4,5,5',6-heptachlorobiphenyl
93	1682.65	174	7	3	2,2',3,3',4,5,6'-heptachlorobiphenyl
94	1692.41	177	7	3	2,2',3,3',4',5,6-heptachlorobiphenyl
95	1700.57	171	7	3	2,2',3,3',4,4',6-heptachlorobiphenyl
96	1703.08	156	6	1	2,3,3',4,4',5-hexachlorobiphenyl
97	1709.56	173	7	3	2,2',3,3',4,5,6-heptachlorobiphenyl
98	1713.95	157	6	1	2,3,3',4,4',5'-hexachlorobiphenyl
		200	8	4	2,2',3,3',4,5',6,6'-octachlorobiphenyl
99	1725.45	172	7	2	2,2',3,3',4,5,5'-heptachlorobiphenyl
		197	8	4	2,2',3,3',4,4',6,6'-octachlorobiphenyl
100	1737.14	180	7	2	2,2',3,4,4',5,5'-heptachlorobiphenyl
101	1741.77	193	7	2	2,3,3',4',5,5',6-heptachlorobiphenyl
102	1748.94	191	7	2	2,3,3',4,4',5',6-heptachlorobiphenyl
103	1756.06	199	8	4	2,2',3,3',4,5,6,6'-octachlorobiphenyl
104	1774.28	169	6	0	3,3',4,4',5,5'-hexachlorobiphenyl
105	1790.56	170	7	2	2,2',3,3',4,4',5-heptachlorobiphenyl
		190	7	2	2,3,3',4,4',5,6-heptachlorobiphenyl
106	1800.60	198	8	3	2,2',3,3',4,5,5',6-octachlorobiphenyl
107	1807.87	201	8	3	2,2',3,3',4,5,5',6'-octachlorobiphenyl
108	1817.30	203	8	3	2,2',3,4,4',5,5',6-octachlorobiphenyl
		196	8	3	2,2',3,3',4,4',5,6'-octachlorobiphenyl
109	1845.47	189	7	1	2,3,3',4,4',5,5'-heptachlorobiphenyl
110	1870.77	195	8	3	2,2',3,3',4,4',5,6-octachlorobiphenyl
111	1883.67	207	9	4	2,2',3,3',4,4',5,6,6'-nonachlorobiphenyl
112	1910.60	194	8	2	2,2',3,3',4,4',5,5'-octachlorobiphenyl
		208	9	4	2,2',3,3',5,5',6,6'-nonachlorobiphenyl
113	1919.37	205	8	2	2,3,3',4,4',5,5',6-octachlorobiphenyl
114	1969.98	(Internal Standard)			octachlorobiphenyl
115	1977.43	206	9	4	2,2',3,3',4,4',5,5',6-nonachlorobiphenyl

Table 2. Charterization of PCB congeners eluted from a column packed with 300 mg of sieved (2–10 μm) AMOCO PX-21 carbon dispersed on 4.5 g of Whatman GF/D filter material

30% Methylene chloride fraction compound (IUPAC number)	100% Toluene fraction compound (IUPAC number)
2,2',3,3',4-PnCB (82)	2,3,3',4'-TCB (56) &/or 2,3,4,4'-TCB (60)*
2,2',3,4,5'-PnCB (87)	3,4,4',5-TCB (81)*
2,2',3',4,5-PnCB (97)	3,3',4,4'-TCB (77)*
2,2',4,4',5-PnCB (99)	2,3,3',4',5-PnCB (107)
2,2',4,5,5'-PnCB (101)	2,3,3',4,4'-PnCB (105)*
2,3,3',4',6-PnCB (110)	2,3,3',4',6-PnCB (110)
2,2',3,3',4,4'-HxCB (128)*	2,3,4,4',5-PnCB (114)*
2,2',3,3',4,5-HxCB (129)	2,3',4,4',5-PnCB (118)*
2,2',3,3',4,5'-HxCB (130)	2',3,4,4',5-PnCB (123)*
2,2',3,3',4,6'-HxCB (132)	2',3,4,5,5'-PnCB (124)
2,2',3,4,4',5-HxCB (137)	2,2',3,4,4',5'-HxCB (138)*
2,2',3,4,4',5'-HxCB (138)*	2,3,3',4,4',5-HxCB (156)*
2,2',3,4',5,5'-HxCB (146)	2,3,3',4,4',5'-HxCB (157)*
2,2',3,4',5',6-HxCB (149)	2,3,3',4,4',6-HxCB (158)*
2,3,3',4,4',6-HxCB (158)*	2,3',4,4',5,5'-HxCB (167)*
2,3,4,4',5,6-HxCB (166)*	3,3',4,4',5,5'-HxCB (169)*
2,2',3,4,4',5,6'-HpCB (182) &/or 2,2',3,4',5,5',6-HpCB (187)	2,2',3,3',4,4',5-HpCB (170)*
2,2',3,4,4',5',6-HpCB (183)	2,2',3,4,4',5,5'-HpCB (180)
2,2',3,3',4,5,5'-HpCB (172) &/or 2,2',3,3',4,4',6,6'-OCB (197)	2,3,3',4,4',5,5'-HpCB (189)
2,2',3,4,4',5,5'-HpCB (180)	2,2',3,3',4,4',5,5'-OCB (194)
2,2',3,3',4,4',5-HpCB (170)*	2,3,3',4,4',5,5',6-OCB (205)
2,3,3',4',5,5',6-HpCB (193)	
2,2',3,4,4',5,5',6-OCB (203) &/or 2,2',3,3',4,4',5,6'-OCB (196)	
2,2',3,3',4,4',5,6-OCB (195) &/or 2,2',3,3',4,5,5',6,6'-NCB (208)	

* AHH Active PCB congeners

Gas Chromatography and Integration

To separate PCB congeners, a fused silica capillary chromatographic column (60 M × 0.25 mm i.d. 0.25 μm film thickness) was used, coated with bonded phase DB5 (J&W Scientific, Inc). To separate AHH-active PCB congeners, we used a more polar, DB1, fused silica capillary chromatographic column (30 M × 0.25 mm i.d. 0.25 μm film thickness) (J&W Scientific, Inc). For either analysis, a 60-cm uncoated fused silica retention gap connected the injector to the analytical column. Hydrogen (linear velocity 32 cm/min) was used as the carrier gas, and nitrogen (15 ml/min) as the detector makeup gas. An IBM 9000 microcomputer, interfaced with the GC, acquired data generated by the electron capture detector (ECD). The data were pre-processed by a software package designed for laboratory data collection (Capillary Applications Program, by IBM Instrument Division, Danbury, CT). Initiation of integration and the GC temperature program were controlled by a Varian Autosampler Model 8000, which also delivered a calibrated amount of sample to the Varian 3700 gas chromatograph injection port.

Chromatographic conditions were similar for all of the analyses: initial temperature, 70°C, programmed at 1°C/min to a final temperature of 255°C; injector temperature (direct inject) 230°C, and detector temperature 320°C. Once the congeners were adequately resolved, the quantitation of PCB congeners could be undertaken.

Peak Identification and Laboratory Data Base

We elected 113 GC peaks (Table 1) of PCB for quantitation and effected calibration by using Aroclors 1242, 1248, 1254, and 1260 in a 1:1:1:1 (w/w/w/w) mixture. A chromatogram of this mixed Aroclor standard is shown in Figure 1 (A–D). The method of peak identification was a retention index system in which *n*-alkyl trichloroacetates were used (Schwartz *et al.* 1983). Four *n*-alkyl trichloroacetates (hexyl, heptyl, octyl, and eicosanyl) were added to the samples before their analysis by GC at a concentration of 0.3 ng/μl. The peak number, retention indices, IUPAC number, and chemical structure for the 113 peaks quantitated in this study are given in Table 1. Once the chromatographic data were collected and preprocessed, the data were organized, by using a BASIC program, into a series of files on hard disk media and transferred off-line to a VAX minicomputer (Digital Equipment Corp.). The data were organized into tree-structured disk files, using specialized laboratory data base management computer programs written in VAX-DSM (Digital Standard MUMPS) for the VAX-DEC (Digital Equipment Corporation) family of computers (Schwartz *et al.* 1984). For data evaluation by principal components analysis, we used SIMCA-3B for PC-DOS microcomputers (Principal Data Components, Columbia, MO).

An example summary report from the PCB analysis of Aroclors and Forster's tern eggs is presented in Table 3. The results and discussion of the polychlorinated dibenzo-*p*-

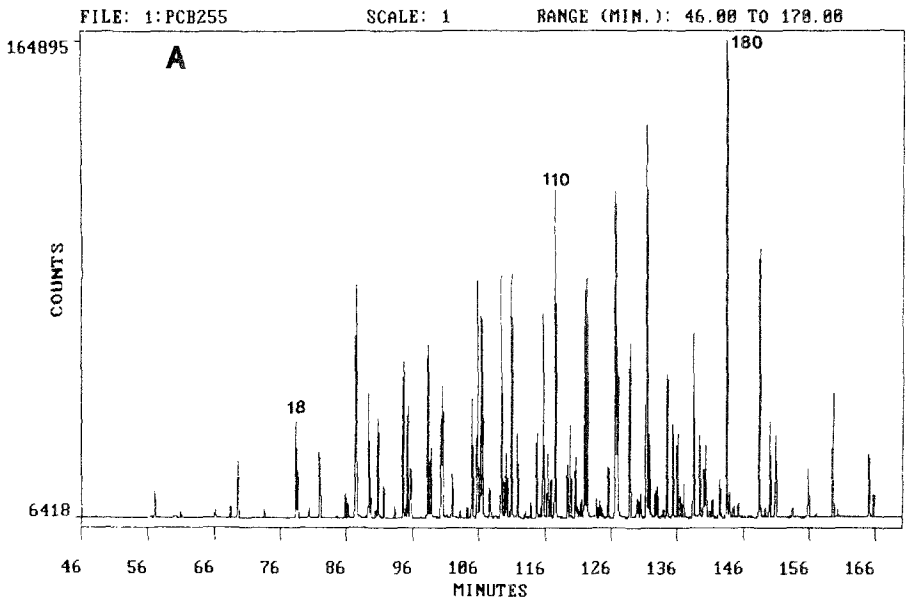

Fig. 1-A. Temperature programmed chromatographic separation of 1:1:1:1 mixture of Aroclor® 1242:1248:1254:1260 on 60-m DB-5 column (see text for chromatographic conditions and Table 1 for peak identity)

dioxins (PCDD), dibenzofurans (PCDF), and non-*ortho*, *ortho*-substituted PCBs concentrations are discussed elsewhere (Kubiak *et al.* 1989).

In addition to the egg samples collected from the Forster's tern colonies a series of Aroclors were analyzed by these techniques to provide a training data set for pattern recognition, and to establish quality control criteria for the data set (Tables 4 and 5).

Principal Component Plots

The PC models are bilinear projection models obtained by decomposing a class data matrix X into a score matrix T (n × F), a loading matrix P (F × p), and a residual matrix E:

$$X = I \cdot x + T \cdot P + E \qquad (1)$$

The objective was to derive models of the isomer and congeners obtained from the data set through a data matrix, X, having *n* objects (40 samples) and *p* variables (113 GC peaks) in which the concentration value of the PCB congener, x_{pn}, could be calculated. The diagonal matrix x is the mean of variables x(p) in all samples. Because we are concerned only with relative differences between samples, the averages, x_{pn}, of each variable are first subtracted from each column in X. This "centering" moves the coordinate system so that the origin is in the center of the data set. The (n × F) score matrix, T, describes the projection of the n sample points onto the F-dimensional hyperplane defined by the (F × p) loading matrix, P. Plotting the columns in T gives a picture of the relations between samples, such as distances, similarities, outliers, etc. Analogously, plots of the rows in the matrix P gives a picture of how the table columns (the variables) are related. Thus, PCA projections extract the information from a data table and present the information graphically. If the residuals, E (or unexplained part of the measurement not modeled), are small in comparison with the variation in X, then the model is a good representation of X.

Normalization and Scaling of Data Prior to Modeling

Two questions concerning pattern recognition are important in deciding what preprocessing of data should be done: What is the composition of the residue? and what is the relative concentration of analytes in each sample?

In approaching the first question, normalization of the concentration data to sum 1 or 100 is recommended, to avoid the influence of absolute concentration. If data are not normalized, the first principal component will strongly correlate with total PCB concentration.

When measured sample data contain a few variables (GC peaks), that are large in relation to the remaining variables and these data are expressed as fractional composition (normalized), the small variable is greatly influenced by small changes in the larger variables. This problem is referred to as closure of the data. Caution is appropriate with respect to closure induced by normalization. Johansson *et al.* (1984) recommended that normalized data be log transformed to further decrease closure. However, closure has not been a significant problem in the capillary column analysis of PCB residues because no single peak dominates the composition of the Aroclors or environmental residues (Stalling *et al.* 1985b). After samples having similar residue profiles have been identified and selected, further peripheral components modeling of non-normalized data can focus on variations in total PCB concentration among samples.

Scaling of Data

Because the outcome of principal components modeling depends on the scaling applied to the data, it is important to consider which scaling methods are appropriate (Derde *et al.* 1982; Sharaf *et al.* 1986). Autoscaling or class scaling are among the most frequently used methods (Schiffman *et al.* 1981; Sharaf *et al.* 1986). Log transformations of data with a

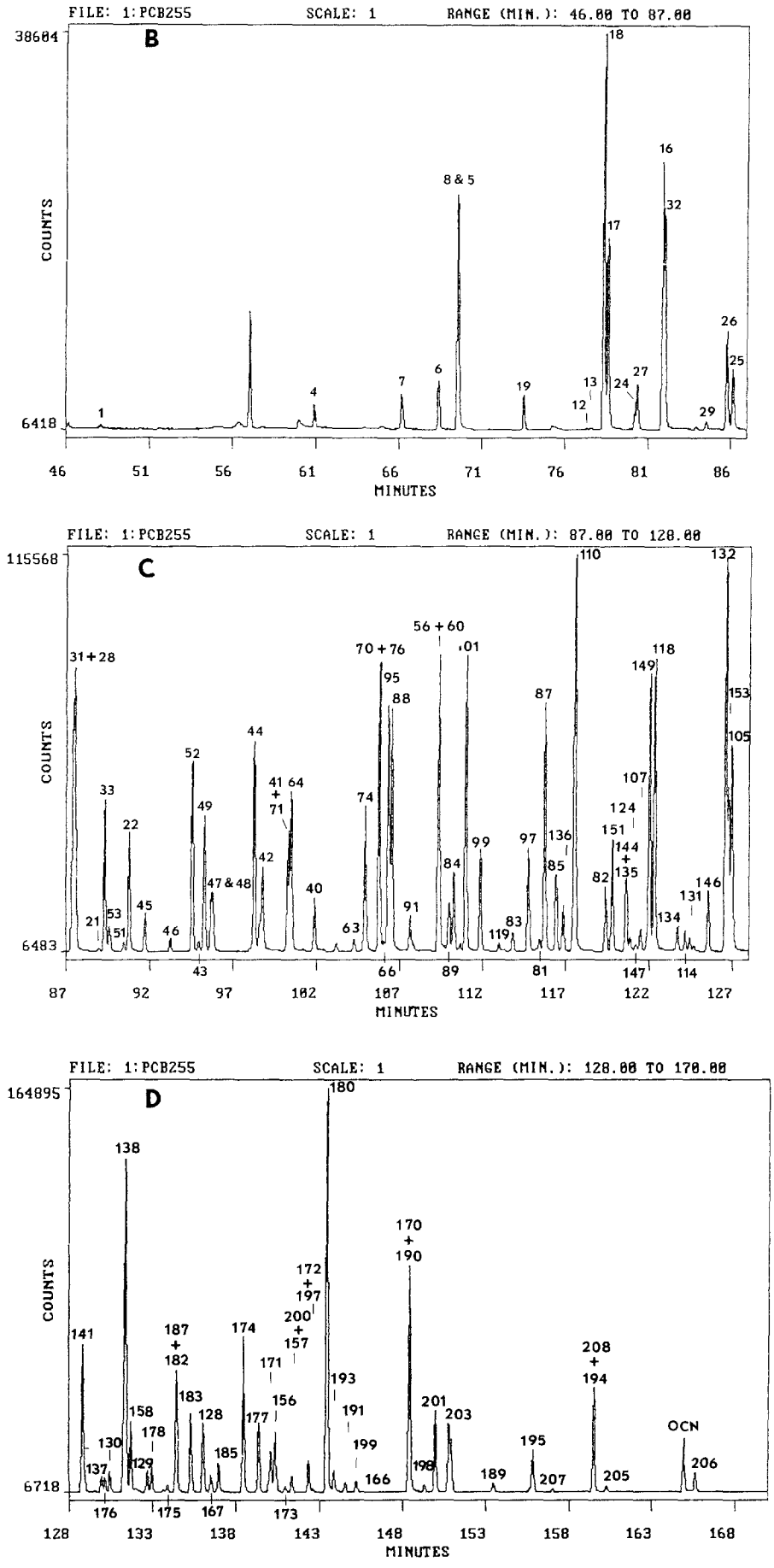

Fig. 1-B,C,D. Segments of Figure 1-A chromatogram (For information concerning peak numbers refer to Table I)

Table 3. Peak number, retention index, IUPAC number, and normalized congener composition in Aroclor® 1242 (A1000), 1248 (A0100), 1254 (A0010), their equal mixture (A1111), and two samples of Forster's Tern eggs (See Table 1 for structure information)

Peak #	RI	IUPAC	A1111	A1000	A0100	A0010	A0001	(25) GF358[a]	(39) PF318
1	720.22	001	0	0	0	0	0	0	0
2	842.64	004 + 010	9.43E-02	0.039	0	0	0	0	0
3	899.70	007	2.50E-03	1.03E-02	0	0	0	0	0
4	923.50	006	4.39E-03	176E-03	0	0	0	0	0
5	936.07	008 + 005	2.18E-02	8.52E-02	9.36E-03	0	0	0	0
6	979.22	019	3.41E	1.16E-02	3.87E-02	0	0	0	0
7	997.31	030 (I.S.)	—	—	—	—	—	—	—
8	1019.69	012	0	0	0	0	0	0	0
9	1022.50	013	0	0	0	0	0	0	0
10	1031.09	018	3.76E-02	0.106	5.15E-02	0	0	0	0
11	1034.05	017	1.58E-02	5.12E-02	1.59E-02	0	0	0	0
12	1051.72	024	9.02E-04	3.26E-03	0	0	0	0	0
13	1053.23	027	1.63E-03	5.57E-03	1.81E-03	0	0	0	0
14	1069.65	016	1.08E-02	3.12E-02	1.06E-02	0	0	0	0
15	1070.84	032	3.23E-03	1.93E-02	5.36E-03	0	0	0	0
16	1098.45	029	0	0	0	0	0	0	0
17	1111.92	026	4.98E-03	1.63E-02	5.77E-03	0	0	0	0
18	1116.04	025	2.40E	8.76E-03	2.02E-03	0	0	0	0
19	1130.10	031 + 028	2.43E-02	8.45E-02	0.101	1.17E-03	0	0	0
20	1146.01	021	0	0	0	0	0	0	0
21	1150.11	033	0	0	0	0	0	0	0
22	1153.31	053	5.25E-03	8.98E-03	1.15E-02	0	0	0	0
23	1162.39	051	1.86E-03	3.80E-03	3.98E-03	0	0	0	0
24	1165.79	022	1.26E-02	0.037	1.51E-02	0	0	0	0
25	1175.56	045	6.04E-03	1.21E-02	0.013	0	0	0	0
26	1192.46	046	2.89E-03	5.70E-03	6.29E-03	0	0	0	0
27	1206.41	052	4.24E-02	4.40E-02	6.69E-02	5.19E-02	4.33E-03	3.62E-02	0
28	1210.99	043	1.40E-03	2.66E-03	2.88E-03	0	0	0	0
29	1214.23	049	2.60E-02	3.86E-02	5.06E-02	1.34E-02	0	2.25E-02	0
30	1219.05	047	6.35E-03	1.04E-02	1.21E-02	2.71E-03	0	1.58E-02	9.75E-03
31	1220.04	048	6.12E-03	1.24E-02	1.54E-02	0	0	0	0
32	1246.95	044	3.06E-02	4.26E-02	5.79E-02	2.14E-02	0	1.90E-02	0
33	1252.52	042	1.79E-02	2.72E-02	3.26E-02	3.28E-03	0	0.011	0
34	1269.27	041 + 071	1.33E-02	0.025	3.02E-02	4.19E-03	0	0	0
35	1271.07	064	1.40E-02	1.80E-02	0.028	5.67E-03	0	1.39E-02	3.89E-03
36	1286.59	040	6.29E-03	1.11E-02	1.32E-02	0	0	0	0
37	1300.45	CL4	0	2.58E-03	2.05E-03	0	0	0	0
38	1312.16	063	1.41E-03	2.34E-03	3.12E-03	0	0	0	0
39	1319.22	074	1.60E-02	2.25E-02	3.25E-02	8.16E-03	0	1.97E-02	6.59E-03
40	1328.17	070 + 076	3.62E-02	4.24E-02	6.62E-02	3.15E-02	0	9.11E-03	0
41	1331.42	066	1.30E-03	0	2.37E-03	1.77E-03	0	0	0
42	1334.23	095	3.21E-02	4.07E-02	7.07E-002	1.05E-02	0	5.31E-02	1.44E-02
43	1336.35	088	2.87E-02	0	1.42E-02	5.88E-02	3.42E-02	1.03E-02	0
44	1348.57	091	5.49E-03	2.63E-03	7.59E-03	9.72E-03	0	0	0
45	1367.03	056 + 060	2.21E-02	3.47E-02	4.67E-02	5.35E-03	0	0	0
46	1373.68	089	6.08E-03	2.37E-03	4.60E-03	1.41E-02	5.33E-03	1.48E-02	7.35E-03
47	1376.79	084	6.64E-03	4.68E-03	1.00E-02	2.02E-02	1.99E-03	0	0
48	1381.03	CL5	6.53E-04	0	1.33E-03	0	0	5.31E-03	0
49	1384.58	101	4.02E-02	8.21E-03	2.13E-02	8.48E-02	4.97E-02	4.76E-02	1.82E-02
50	1394.36	099	1.41E-02	5.50E-03	1.44E-02	3.26E-02	0	4.08E-02	3.93E-02
51	1406.31	119	7.04E-04	0	0	1.55E-03	0	0	0
52	1415.48	083	2.22E-02	0	2.55E-03	4.86E-03	0	0	0
53	1425.50	097	1.17E-02	5.03E-03	1.23E-02	3.93E-02	1.51E-03	1.26E-02	0
54	1432.80	081	1.70E-03	0	1.93E-03	4.01E-03	0	0	0
55	1436.12	087	1.98E-02	6.35E-03	1.56E-02	4.75E-02	7.26E-03	1.44E-02	0
56	1443.54	085	5.82E-03	3.52E-03	7.98E-03	1.13E-02	0	1.97E-02	1.34E-02
57	1448.43	136	0	0	0	0	0	0	0
58	1454.09	077	0	0	0	0	0	0	0
59	1455.72	110	2.45E-02	9.83E-03	2.49E-02	6.20E-02	1.95E-02	4.43E-02	0.022
60	1476.03	082	4.53E-03	2.93E-03	6.56E-03	8.67E-03	0	8.30E-03	0
61	1480.50	151	1.23E-02	0	0	9.27E-03	4.92E-02	5.97E-03	0
62	1489.34	144 + 135	8.00E-03	0	0	1.00E-02	2.64E-02	7.12E-03	0

Table 3. (cont'd)

Peak #	RI	IUPAC	A1111	A1000	A0100	A0010	A0001	(25) GF358[a]	(39) PF318
63	1492.14	124	1.05E-03	0	8.74E-04	3.07E-03	0	0	0
64	1495.94	147	7.21E-04	0	0	2.25E-03	0	0	0
65	1498.75	107	1.93E-03	0	1.91E-03	4.92E-03	0	0	0
66	1503.08	123	0	0	0	0	0	0	0
67	1504.31	149	3.57E-02	0	4.04E-03	4.32E-02	6.99E-02	2.96E-02	2.50E-02
68	1507.72	118	0.028	6.28E-03	1.87E-02	7.32E-02	8.79E-03	4.95E-02	4.09E-02
69	1522.78	134	2.72E-03	0	0	5.51E-03	6.00E-03	0	0
70	1527.46	114	9.96E-04	0	1.34E-03	2.14E-03	0	2.99E-03	0
71	1530.70	131	1.30E-03	0	0	2.87E-03	2.56E-03	0	5.14E-03
72	1533.36	122	4.97E-04	0	0	0	0	0	0
73	1542.55	146	5.51E-03	0	0	8.02E-03	1.69E-02	1.86E-02	3.25E-02
74	1553.92	132	4.26E-02	0	5.77E-03	5.21E-02	0.123	9.63E-02	0.166
75	1555.62	153	3.62E-02	0	0	2.54E-02	0	0	0
76	1557.70	105	1.17E-02	4.44E-03	1.24E-02	2.74E-02	0	2.36E-02	1.74E-02
77	1577.31	141	0.017	0	1.61E-03	1.58E-02	4.44E-02	8.43E-03	0
78	1589.79	137	1.86E-03	0	0	7.30E-03	6.15E-04	5.58E-03	7.09E-03
79	1592.57	176	2.14E-03	0	0	0	0/010	0	0
80	1595.79	130	2.39E-03	0	0	5.71E-03	4.26E-03	7.31E-03	1.01E-02
81	1605.85	138	2.97E-02	6.44E-04	3.86E-03	0.059	5.23E-02	7.36E-02	0.125
82	1609.35	158	4.45E-03	0	5.03E-04	8.40E-03	1.17E-02	6.00E-03	9.58E-03
83	1619.91	129	2.17E-03	0	0	5.85E-03	1.93E-03	0	0
84	1621.54	126	0	0	0	0	0	0	0
85	1622.87	178	2.54E-03	0	0	0	1.17E-02	3.77E-03	5.28E-03
86	1633.03	175	6.74E-04	0	0	0	3.23E-03	4.34E-03	4.18E-03
87	1639.00	187 + 182	1.64E-02	0	1.34E-03	3.56E-03	4.58E-02	3.33E-02	6.34E-02
88	1648.27	183	8.30E-03	0	4.26E-04	2.64E-03	3.92E-02	1.15E-02	2.57E-02
89	1656.37	128	6.38E-03	0	1.21E-03	1.66E-02	8.36E-03	1.51E-02	2.23E-02
90	1661.23	167	1.55E-03	0	0	3.54E-03	3.12E-03	5.17E-03	8.08E-03
91	1666.72	185	1.63E-03	0	0	0	7.71E-03	0	0
92	1682.00	174	1.32E-02	0	1.41E-03	4.69E-03	3.76E-02	9.47E-03	1.41E-02
93	1692.37	177	7.64E-03	0	0	2.76E-03	3.57E-02	8.09E-03	1.75E-02
94	1700.49	171	3.82E-03	0	0	2.01E-03	1.50E-03	6.45E-03	1.30E-02
95	1702.98	156	3.85E-03	0	7.26E-04	8.90E-03	7.35E-03	8.50E-03	1.45E-02
96	1709.63	173	4.47E-04	0	0	0	1.06E-03	0	0
97	1713.94	200 + 157	1.29E-03	0	0	2.29E-03	3.17E-03	0	841E-03
98	1724.80	172 + 197	1.99E-03	0	0	9.69E-04	9.19E-03	5.29E-03	0
99	1736.91	180	2.06E-02	0	2.04E-03	8.68E-03	7.77E-02	5.21E-02	9.75E-02
100	1741.68	193	1.41E-03	0	0	0	6.68E-03	2.85E-03	4.58E-03
101	1748.77	191	6.25E-04	0	0	0	3.87E-03	0	0
102	1748.94	199	0	0	0	0	0	0	0
103	1774.24	169	0	0	0	0	0	0	0
104	1790.53	170 + 190	1.58E-02	0	1.11E-03	7.28E-03	4.04E-02	2.06E-02	3.97E-02
105	1800.51	198	5.58E-04	0	0	0	2.45E-03	2.79E-03	3.36E-03
106	1807.58	201	5.91E-03	0	7.86E-04	0	3.09E-02	1.26E-02	0.018
107	1816.52	203 + 196	4.37E-03	0	8.00E-04	0	1.21E-02	1.46E-02	2.35E-02
108	1845.17	189	6.37E-04	0	0	0	2.28E-03	0	4.11E-03
109	1870.78	195	2.49E-03	0	0	2.48E-03	1.28E-02	3.50E-03	5.75E-03
110	1883.80	207	2.78E-04	0	0	0	1.19E-03	0	0
111	1910.67	208 + 194	5.68E-03	0	5.89E-04	0	2.20E-02	9.70E-03	1.72E-02
112	1918.97	205	4.36E-04	0	0	0	1.65E-03	0	4.72E-03
113	1969.82	OCN	—	—	—	—	—	—	—
114	1997.63	206	9.80E-04	0	0	0	4.70E-03	2.66E-03	5.10E-03
Total PCB concentration (μg/g)			7.90	2.43	3.07	2.63	4.13	14.1	2.90

[a] Eggs collected from colony on Green Bay, WI in 1983 (SIMCA number) Sample I.D.
[b] Eggs collected from colony on Lake Poygan, WI in 1983 (SIMCA number) Sample I.D.

zero offset is another method that is effective in unsupervised pattern recognition (Kvalheim 1985).

Autoscaling results from setting the individual means of each variable selected equal to zero and the scaling of the variance of each variable to unit variance. The outcome of this method of scaling is that each variable assumes equal importance in the model and measurements of dissimilar type may be compared. This method is recommended for

Table 4. Sample identification and total PCB residues measured in eggs from Forster's Terns collected in 1983 from colonies on Lake Poygan and Green Bay, Wisconsin

Green Bay (G)	PCB (μg/g)	Lake Poygan (P)	PCB (μg/g)
25 GF358	15.3	33[a] PF320	8.03
26 GF359	25.6	34 PF319	4.88
27 GF356	6.6	36 PF315[b]	5.71
28 GF357	21.7	37 PF316	4.56
29 GF344	24.3	38 PF317	2.65
30 GF361	28.0	39 PF318	2.90
31 GF360	26.8		
32 GF344	25.0		
35 GF221[c]	26.2		
40 GF358	10.2		
Mean (SD)	21.0 (7.56)		4.79 (1.98)

[a] Sample number for principal components sample score plots
[b] Sample excluded from Lake Poygan principal components model First letter designates sample origin, code following is laboratory i.d. number
[c] Sample collected in 1982 from colony on Green Bay

Table 5. Aroclors® and their mixtures analyzed for principal components modeling

Aroclor sample identification	
1[a] A1111 1:1:1:1[b]	13 A1100 1:1:0:0
2 A1000	14 A1010 1:0:1:0
3 A0100	15 A1001 1:0:0:1
4 A0010	16 A0110 0:1:1:0
5 A0001	17 A0101 0:1:0:1
6 A1111 1:1:1:1	18 A0011 0:0:1:1
7 A0001	19 A0011 0:0:1:3
8 A0001	20 A0011 0:0:3:1
9 A0010	21 A1110 1:1:1:0
10 A0010	22 A1011 1:0:1:1
11 A1111 1:1:1:1	23 A1101 1:1:0:1
12 A1000	24 A0111 0:1:1:1

[a] Sample number for principal components sample score plots
[b] A1111 1:1:1:1. A one (1) following the letter A designates the addition of one or more Aroclor in the order Aroclors® 1242:1248:1254:1260 and the numbers separated by colons designate the weight ratio(s)

GC analysis for classification purposes (Kowalski 1973; Massart *et al.* 1978). In attempts to describe PCB residue profiles by using normalized data, however, it has not been necessary to autoscale the data except in classification problems.

Chemometric Analysis of Forster's Tern Eggs

The analysis and interpretation of PCB residues in Forster's terns from two locations in Wisconsin provide a good example of using a chemometrics approach to data analysis by pattern recognition.

Eggs were collected in 1983 from Forster's terns from colonies on Green Bay and Lake Poygan, WI, by U.S. Fish and Wildlife Regional personnel as part of a biological survey relating to environmental contaminants. Birds from Lake Poygan served as a control site for the survey. In addition, one egg sample collected from a Green Bay colony in 1982 was available for analysis.

The total PCB concentration reported (Table 4) was obtained by summing the concentration of each congener before the measured concentrations reported were normalized. The average total concentration of PCBs in the eggs from Green Bay was 21.0 mg/kg (7.56 s.d., n = 6), as compared with 4.79 mg/kg (1.98 s.d., n = 10) in eggs from Lake Poygan.

The Aroclor composition and residues detected in the egg samples from Lake Poygan and Green Bay collections reported in Table 3 illustrates the difficulty of evaluating residue data when large numbers of variables (congeners) are measured. The data obtained from the analysis of the Forster's terns, Aroclors, and Aroclor mixtures form a matrix of 40 samples in which 113 congener constituents have been measured, which in addition to the total concentration for each sample, represent 4,520 observations. Although tabular data may be comprehensible, one can gain only a limited perspective of sample similarity from the traditional approach of examining one variable at a time. Clearly, a better approach to data analysis is needed. Principal components analysis offers such an approach to multivariate data. Several questions were appropriately addressed through principal components modeling of these analytical results: the relations among Aroclors and PCB residues in egg; the determination of residue profiles of PCB relative to that of commercial Aroclor mixtures and the dependence of location on residue profile.

Graphical Representation of PCB Congeners from PCA

Stalling *et al.* (1985b) described the geometric relations of Aroclors and their mixtures by examining principal component models of Aroclors and their mixtures. From the data set used to describe Tern egg PCB residue profiles, a three-term PCA model was found to be significant and described more than 84% of the total variance for this data set (Table 6).

General information about sample similarity and relations among Aroclors, their mixtures, and Tern egg samples can be obtained from plotting three-dimensional projections of sample scores using a BASIC program (available from the authors on request) (Figure 2). The relations in this plot are more readily comprehended when examined in color (plotted points have contrasting colors designating each class of samples). From an examination of the 3-D projection of Aroclors and egg samples (Figure 2), it is concluded that the residues in eggs have profiles different from those that can be represented by Aroclors or their mixtures, and that the residue profiles in eggs from Green Bay and Lake Poygan colonies may also differ. Although it is difficult to visualize Figure 2 in the two dimensional plane of the paper, the Lake Poygan cluster of samples clearly lies in front of, and is well separated from, the Green Bay cluster. Three

Table 6. Summary of principal components modeling of PCBs in Aroclors® and Forster's Tern eggs from Green Bay and Lake Poygan, Wisconsin

Class description	N	Total variance	Variance explained (Cumulative %)	PC1	PC2	PC3
Aroclors®	23	1.01×10^{-4}	66.4	89.4		95.5
All samples	39	1.22×10^{-4}	50.8	67.7		84.0
Eggs						
Lake Poygan and Green Bay	16	5.94×10^{-5}	55.5	81.5		ns[a]
Lake Poygan	5	6.40×10^{-5}	72.6	89.6		ns
Green Bay	9	1.31×10^{-5}	39.0	68.9		ns

[a] ns = cross validation not significant

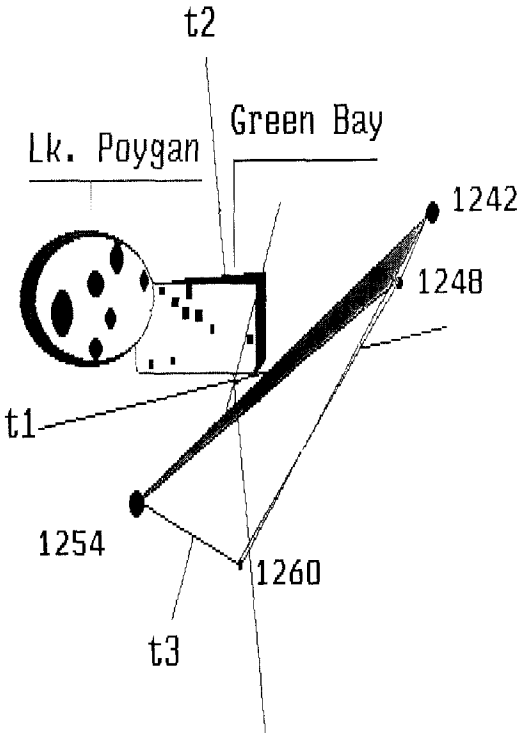

Fig. 2. Three dimensional principal components score plot for Aroclors, their mixtures, and PCB residues in Forster's Tern eggs from Lake Poygan and Green Bay, Wisconsin

dimensional rotation on a computer CRT clearly shows spacial relationships. The clustering of samples from Green Bay and Lake Poygan can be examined quantitatively through classification studies.

The similarity concept for this study is based on two fundamentals and can be quantified in situations where two conditions are met: (1) measurements of the same type have been made on a number of systems (ie: the same PCB isomers have been measured in all samples), and (2) the samples have been ordered into classes each of which contains only similar samples (ie: the classes in our study were the Aroclor standards, bird eggs from Green Bay and bird eggs from Lake Poygan). The objective of classification is to derive an appropriate description of the data structure within each class, in terms of a quantitative model. Class integrity is then tested by using PCA to determine whether samples have been assigned to the correct class. The similarity of samples within the class can be assessed by the proximity of samples to each other in plots derived from principal components models.

The statistical technique of cross-validation was used to determine that two principal components were statistically significant for each sample class (Green Bay and Lake Poygan eggs) (Wold, S. 1978, Wold, S. 1989) (Table 6). In addition to classifying the sample residues as derived from Green Bay or Lake Poygan eggs, Aroclor and Aroclor mixtures were evaluated with respect to each class model to determine residue composition similarity.

If there is a high degree of probability that Aroclors or mixtures of Aroclors can be classified as members of either or both of the class models for Green Bay and Lake Poygan, the question becomes, What mixture of Aroclors best describes the samples in the respective models? If the Aroclors or their mixtures cannot be classified as members of either class model within a reasonable degree of probability (99%) the question need not be posed because the samples are not judged to be Aroclors and cannot be described by any combination of Aroclors. It would be inappropriate to report the PCBs in these samples in terms of relative Aroclors composition.

The results of the classification study are shown graphically in Figure 3. To prepare the plot, the distance of each sample from the class model selected is calculated in units of class standard deviation. The X and Y coordinates of the sample in the plot correspond to the distances from the Lake Poygan model and Green Bay respectively. The distance corresponding to the 0.01 level of certainty derived from F-statistics for both classes are plotted as horizontal and vertical dashed lines (Wold *et al.* 1984b). All samples show a high degree of membership in one of the two classes, with exceptions noted below, indicating that the data are well described by the class models. Residue profiles in the two sets of eggs differed from each other and none of the Aroclors or their mixtures fell within the class of samples originating from either Lake Poygan or Green Bay—confirming the conjecture made from the clustering of samples in Figure 2. One egg sample (Sample 35, collected in 1982), falls outside the class model of both egg groups and is considered an outlier for either class. Because the 1982 sample differed markedly from the 1983 samples, we are curious about year-to-

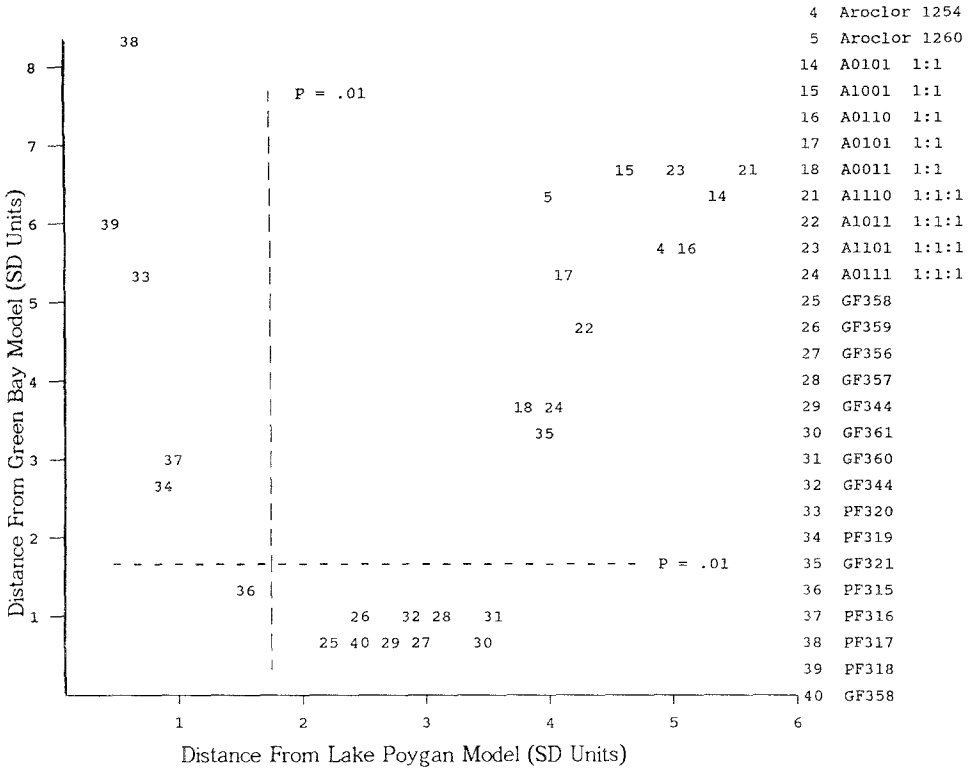

Fig. 3. Cooman's plot of sample distances from Lake Poygan (X-axis) (Sample #, LF---), and Green Bay (Y-axis) (Sample #, GF---), WI Green Bay and Lake Poygan principal components class models for congener-specific PCB residues (Horizontal and vertical dashed-lines represent p = .01 confidence limits of class models)

year changes in residue profiles. Another egg (Sample 36, PF315) falls within both classes at the p = .01 level. One explanation for the classification of Sample 35 in both classes is that this egg may be from a bird originally a member of the Green Bay colony which moved to Lake Poygan.

The question of whether PCB profiles can be accurately represented by an Aroclor or by Aroclor mixtures is approached quantitatively using the Aroclors as a training (or reference) set and then classifying the samples to determine if they can be considered as a member of the Aroclor class. This plot (not shown) revealed that the sample collected in 1982 (Sample 35 in Figure 3) was similar to the Aroclor class, but that the samples collected in 1983 were not.

The total PCB residues in the two groups differed by four-fold. However, the difference in total concentration is not directly related to differences detected by the principal components modeling because the samples were normalized before modeling. The samples from Lake Poygan do not cluster tightly because the compositions of the residues vary. A review of the chromatograms revealed that ratios of late and middle peaks varied, perhaps because there is more than one source of PCB contamination in the diet of Lake Poygan birds. A more critical study of the relations between chlorine substitution patterns should be undertaken to explore the question of whether the differences in profiles between the two colonies reflect variation in dietary contamination or are related to the capacity of the birds to metabolize PCBs before they are deposited in the egg.

Principal Component Plots of AHH-Active PCB Congeners

The question of overall sample similarity of AHH-active congener concentrations from the bird eggs and Aroclors is examined by a global PCA model of those congeners. Two principal components terms were found to be significant and accounted for 94% of the variance. From an examination of Figure 4, it is concluded that the residues in eggs form distinct clusters separated according to location, Green Bay and Lake Poygan. Residue profiles different from those that can be represented by Aroclors with the possible exception of Aroclor 1254 which is located in the Green Bay tern egg cluster. However, a great deal of the variance in the principal component model is due to the Aroclors outside the two clusters. A second model, shown in Figure 5, explained 96% of the variance after eliminating the Aroclors with the exception of 1254. It can be concluded from Figure 5 that the residue profiles in eggs from Green Bay and Lake Poygan colonies differ from each other and also differ from the AHH-active congeners of Aroclor 1254. These interpretations can be examined more quantitatively by classification studies similar to those described above. PCA models were calculated for Lake Poygan and Green Bay samples. Both of the egg classes were described by models with two principal components terms describing 83% and 82% of the variance respectively. Aroclors and samples were fitted to each of the two egg class models and classified as a member or not a member of each class. As described above, the dis-

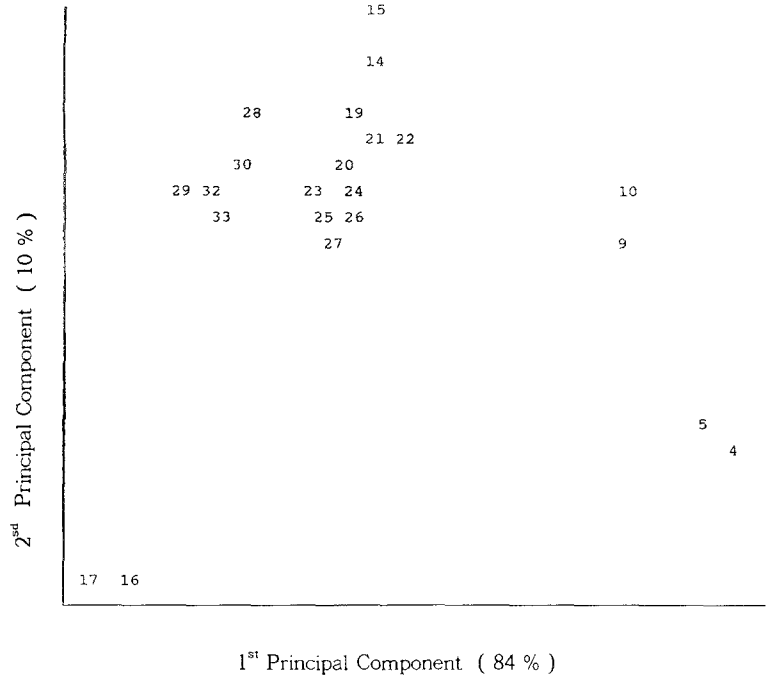

Fig. 4. Plot of sample scores from two-term principal components model of AHH-active PCB residues in Aroclors and Forster's Tern eggs from Green Bay (Sample #, GF---) and Lake Poygan (Sample #, PF---)

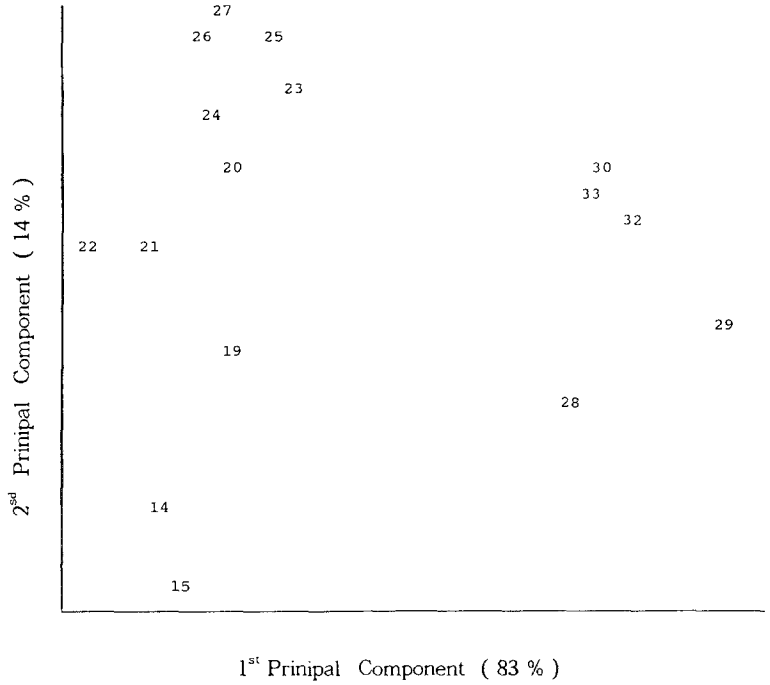

Fig. 5. Plot of sample scores from two-term principal components model of AHH-active PCB residues in Aroclor® 1254 and Forster's Tern eggs from Green Bay (Sample #, GF---) and Lake Poygan (Sample #, PF---)

tance of each sample from the class model selected is calculated in units of class standard deviation. One then makes a plot, using each sample's distance from two classes as coordinates. In Figure 6, the X and Y coordinates of the sample in the plot correspond to the distances from the Lake Poygan and Green Bay respectively. The distance corresponding to the 0.01 level of certainty derived from F-statistics for both classes are plotted as horizontal and vertical dashed lines. From this plot, we conclude that profiles in the two sets of eggs differed from each other, and that Aroclor 1254 could not be described by either model derived from the AHH-active residues found in tern eggs from Lake

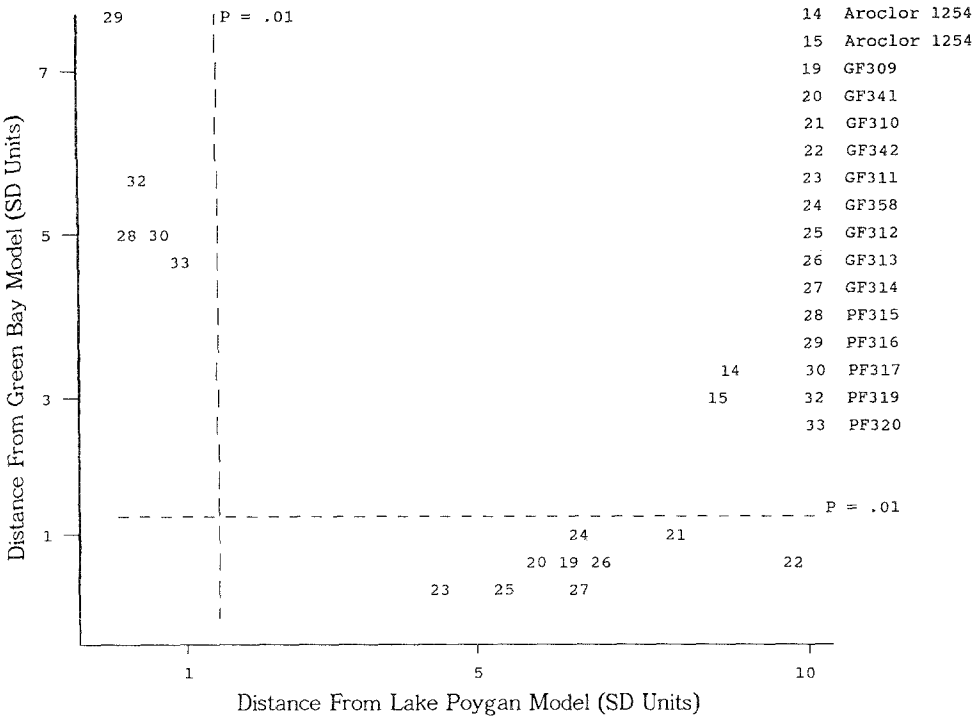

Fig. 6. Cooman's plot of sample distances from Lake Poygan (X-axis) (Sample #, LF---) and Green Bay (Y-axis) (Sample #, GF---), WI Green Bay and Lake Poygan principal complements class models for AHH-active PCB Residues (Horizontal and vertical dashed-lines represent p = .01 confidence limits of class models)

Poygan or Green Bay—confirming the conjecture made from the clustering of samples in Figure 5. These results parallel those of the congener specific procedure. Not all of the samples used in the congener specific procedure, described above, were available for AHH congener analysis. One sample missing from this data set is the tern egg collected in 1983 (Figure 3, Sample 35). In Figure 3, it was speculated that Sample 36 (PF315) was classified as a member of both Green Bay class and Lake Poygan class, because the egg may have come from a bird originally a member of the Green Bay colony which moved to Lake Poygan. Figure 6 clearly shows this egg (Sample 28, PF315 in this AHH congener data set) to be a member of the Lake Poygan colony. The discrepancy between the conclusions drawn for this sample can be explained by the loss of information when the number of variables measured is reduced, a phenomenon discussed previously (Stalling *et al.* 1987).

Conclusions

The task of obtaining a comparable perspective of what relations exist in these data by examining the concentration of individual constituents is formidable. However, through the application of pattern recognition by SIMCA, profiles of PCBs have been characterized in a typical environmental situation and show that residues in bird eggs of two colonies differ in composition. Furthermore, residues should not be reported as Aroclor equivalents in either overall congener composition of AHH active residues. Because the residue profiles cannot accurately be represented as Aroclors, these residues should be reported in terms of total PCB concentration or by combining the isomers into their respective Cl_{1-10} congener sums.

Acknowledgments. The authors wish to thank T. J. Kubiak, for supplying the Forster's tern eggs and many hours of useful discussion. We also wish to thank Jimmy D. Petty, Paul H. Peterman, and William Dunn, III for substantive suggestions on an earlier version of this manuscript. Kevin P. Feltz spent many hours reviewing chromatograms and verifying peak identities, his effort is deeply appreciated.

References

Albano C, Dunn WJ III, Edlund U, Esbensen K, Hellberg S, Johansson E, Lindberg W, Sjöström M (1978) Four levels of pattern recognition. Anal Chim Acta Compu Tech Optim 103:429–443

Albro PW, Corbett JT, Schroeder JL (1981) Quantitative characterization of polychlorinated biphenyl mixtures (Aroclor® 1248, 1254, and 1260) by gas chromatography using capillary columns. J Chromatogr 205:103–111

Ballschmiter K, Zell MZ (1980) Analysis of polychlorinated biphenyls (PCB) by glass capillary gas chromatography. Composition of technical Aroclor- and Clophen®-PCB Mixtures. Z Anal Chem 320:20–31

Bandiera S, Sawyer T, Romkes M, Zmudzka B, Safe L, Mason G, Keys B, Safe S (1984) Polychlorinated dibenzofurans (PCDFs): Effects of structure on binding to the 2,3,7,8-TCDD cytosol receptor protein. Toxicology 32:131–144

Brinkman UA, Th de Kok A (1980) Production, properties and usage. In: Kimbrough RD (ed) Topics in environmental health: Vol. 4. Halogenated biphenyls, terphenyls, naphthalenes, dibenzodioxins and related products, North-Holland, NY, pp 2–7

Bush B, Connor S, Snow J (1982) Glass capillary gas chromatography for sensitive accurate polychlorinated biphenyl analysis. J Assoc Offic Anal Chem 65:555–566

Derde MP, Coomans D, Massart DL (1982) Effect of scaling on class modeling with the SIMCA method. Anal Chim Acta 141:187–192

Duinker JC, Hillebrand JIJ, Palmark KH, Wilhemsen S (1980) An evaluation of existing methods for quantitation of polychlorinated biphenyls in environmental samples and suggestions for an improved method based on measurement of individual components. Bull Environ Contam Toxicol 25:956–964

Dunn WJ III, Stalling DL, Schwartz TR, Hogan JW, Petty JD (1984) Pattern recognition for classification and determination of polychlorinated biphenyls in environmental samples. Anal Chem 56:1308–1313

Environmental Protection Agency (USEPA) (1979) Appendix III. Example quality assurance and quality control procedures for organic priority pollutants. Federal Register 44:69540–69559 (December 3, 1979)

Erikson MD (1986) Analytical Chemistry of PCBs. Ann Arbor Science, Butterworth Publishers, Stoneham, MA, 508 pp

Goldstein JA (1980) Structure-activity relationships for the biochemical effects and the relationship to toxicity: In: Kimbrough RD (ed) Topics in environmental health: Vol. 4. Halogenated biphenyls, terphenyls, naphthalenes, dibenzodioxins and related products, North-Holland, NY, Chapter 6

Jensen S (1966) Report on a new chemical hazard. New Sci 32:612

Johansson E, Sjöström K, Wold S (1984) Minimizing effects of closure on analytical data. Anal Chem 56:1685–1688

Kowalski B (1973) Pattern recognition II. Linear and nonlinear methods for displaying chemical data. J Amer Chem Soc 95:686–693

Kubiak TJ, Harris HJ, Smith LM, Schwartz TR, Stalling DL, Trick JA, Sileo L, Docherty DE, Erdman TC (1989) Microcontaminants and reproductive impairment of the Forster's tern on Green Bay, Lake Michigan-1983. Arch Environ Contam Toxicol 18:706–727

Kvalheim OM (1985) Scaling of analytical data. Anal Chi Acta 177:71–79

Luotamo M, Aitio A, Wold S (1988) Serum polychlorinated biphenyls: Quantitation and Identification of source exposure by the SIMCA pattern recognition method. Chemometrics and Intelligent Laboratory Systems, 4 (2):171–181

McConnell EE (1980) Acute and chronic toxicity, carcinogenesis, reproduction, teratogenesis and mutagenesis in animals. In: Kimbrough RD (ed) Topics in environmental health: Vol. 4. Halogenated biphenyls, terphenyls, naphthalenes, dibenzodioxins and related products, North-Holland, NY, Chapter 5

Massart DL, Dijskstra A, Kaufman L (1978) Evaluation and Optimization of Laboratory Methods and Analytical Procedures. Elsevier, Amsterdam. 596 pp

Mullin MD, Pochini CM, McCrindle S, Romkes M, Safe SH, Safe L (1984) High resolution PCB analysis: Synthesis and chromatographic properties of all 209 PCB congeners. Environ Sci Technol 18:468–476

Onuska FJ, Mudroch A, Davies SJ (1985) Application of chemometrics in homologue-specific analysis of polychlorinated biphenyls. J High Res Chromatog and Chromatog Comm 8:748–754

Polland A, Knutson JC (1982) 2,3,7,8-Tetrachlorodibenzo-*p*-dioxin and related halogenated aromatic hydrocarbons: Examination of the mechanism of toxicity. Annu Rev Pharmacol Toxicol 22:517–554

Rappe C, Buser HR (1980) Chemical Properties and Analytical Methods. In: Kimbrough RD (ed) Topics in environmental health: Vol. 4. Halogenated biphenyls, terphenyls, naphthalenes, dibenzodioxins and related products, North-Holland, NY, pp 41–47

Safe S, Parkinson A, Robertson L, Crockerline R, Safe L, Bandiera S, Okey A (1982) PCBs as AHH inducers. In: Hutzinger, O, Frei, RW, Merian, E, Pocchiari, F, (eds) Chlorinated Dioxins and Related Compounds. Pergamon Press: New York, 1982, pp 383–392

Safe S, Bandiera S, Sawyer T, Robertson L, Safe L, Parkinson A, Thomas PE, Ryan DE, Reik LM, Levin W, Denomme MA, Fujita T (1985) In: Environmental Health Perspectives, US Dept. of Health and Human Services, Publication no. (NIH) 85-218; May 1985, 60, 47–56

Schiffman SS, Reynolds ML, Young FW (1981) Introduction to Multidimensional Scaling: Theory, Methods, and Applications. Academic Press, New York

Smith LM (1981) Carbon Dispersed in Glass Fibers as an Adsorbent for Contaminant Enrichment and Fractionation. Anal Chem 53:2152–2154

Smith LM, Schwartz TR, Kubiak TJ (1986) Occurrence of AHH-Active Polychlorinated Biphenyls, Dibenzodioxins and Dibenzofurans in Lake Michigan Sediment and Biota. The Question of Their Relative Toxicological Significance. Meeting of the American Chemical Society, Division of Environmental Chemistry, Paper 24, Anaheim, CA

Schwartz TR, Smith LM (1987) Determination of AHH-Active PCB Isomers in PCB Mixtures and Environmental Samples. Meeting of the American Chemical Society, Division of Environmental Chemistry, Paper 76, New Orleans, LA

Schwartz TR, Lehmann RG (1982) Determination of Polychlorinated Biphenyls in Plant Tissue. Bull Environ Contam Toxicol 28:723–727

Schwartz TR, Campbell RD, Stalling DL, Little RL, Petty JD, Hogan JW, Kaiser EM (1984) Laboratory data base for isomer-specific determination of polychlorinated biphenyls. Anal Chem 56:1303–1308

Schwartz TR, Petty JD, Kaiser EM (1983) Preparation of *n*-Alkyl Trichloroacetates and Their Use as Retention Index Standards in Gas Chromatography. Anal Chem 56, 1839

Schwartz TR, Stalling DL, Rice CL (1986) Are polychlorinated biphenyl residues adequately described by Aroclor mixture equivalences? Isomer-specific principal components analysis of such residues in fish and turtles. Environ Sci Technol 21:72–76

Sharaf MA, Illman DL, Kowalski BR (1986) Exploratory Data Analysis. In: Chemometrics, John Wiley and Sons, New York, pp 179–296

Sjöström M, Wold S, Lindberg W, Persson JA, Martens H (1983) A multivariate calibration problem in analytical chemistry solved by partial least-squares models in latent variables. Anal Chim Acta 150:61–70

Stalling DL, Tindle RC, Johnson JL (1972) Cleanup of pesticide and polychlorinated biphenyl residues in fish extracts by gel permeation chromatography. J Assoc Offic Anal Chem 55:32–38

Stalling DL, Dunn WJ III, Schwartz TR, Hogan JW, Petty JD, Johansson E, Wold S (1985a) Application of soft independent method of class analogy (SIMCA) in isomer specific analysis of polychlorinated biphenyls. In: Kurtz DA (ed) Trace residue analysis, chemometric estimations of sampling, amount and error. ACS Symposium Series No. 284, American Chemical Society, Washington, CD, pp 195–234

Stalling DL, Schwartz TR, Dunn WJ III, Petty JD (1985b) Soft Independent Methods of Class Analogy: Use in characterizing complex mixtures and environmental residues of polychlorinated biphenyls. In: Breen JJ, Robinson PE (eds) Environmental applications of chemometrics. ACS Symposium Series, No. 292, American Chemical Society, Washington, DC, pp 1–15

Stalling DL, Schwartz TR, Dunn WJ III, Wold S (1987) Classification of polychlorinated biphenyl residues: Isomers vs homologue concentrations in modeling Aroclors and polychlorinated biphenyl residues. Anal Chem 59:1853–1859

Webb RG, McCall AC (1973) Quantitative PCB standards for electron capture gas chromatography. J Chromatogr Sci 11:366–373

Wold S, Sjöström M (1977) SIMCA: A method for analyzing chemical data in terms of similarity and analogy. In: Kowalski BR (ed) Chemometrics, theory and application. ACS Symposium

Series No. 52, American Chemical Society, Washington, DC, pp 243–282

Wold S, Albano C, Dunn WJ III, Edlund U, Esbensen K, Helberg S, Johansson E, Lindberg W, Sjöström M (1984a) Modeling data tables by principal components and pls: class patterns and quantitative predictive relations. Analusis 12:477–485

Wold S, Albano C, Dunn WJ III, Edlund U, Geladi P, Hellberg S, Johansson E, Lindberg W, Sjöström M (1984b) Multivariate data analysis in chemistry. In: Chemometrics mathematics and statistics in chemistry, D. Reidel, Dordrecht, Holland, pp 17–95

Wold S (1978) Cross validatory estimation of the number of principal components in factor and principal component models. Technometrics 20, 397

———(1982) The analysis of multivariate chemical data using SIMCA and MACUP. Kemia-Kemi 9:401–405

———(1989) Multivariate data analysis: Converting chemical data tables to plots. Intelligent Instruments and Computers 7(5), 197

Manuscript received August 10, 1989 and in revised form May 11, 1990.