Sampling design for fish contaminant monitoring program in Lake Michigan. Report 140 [1987]

Staggs, Michael D.
Madison, Wisconsin: Dept. of Natural Resources, [1987]

DEPARTMENT OF NATURAL RESOURCES
RESEARCH
REPORT 140
FEBRUARY 1987

SAMPLING DESIGN FOR FISH CONTAMINANT MONITORING PROGRAM IN LAKE MICHIGAN

By
Michael Staggs
Bureau of Research*, Madison

ABSTRACT

An evaluation of several candidate sampling designs indicated that the mean contaminant level in sampled fish was the best choice for reporting polychlorinated biphenyl (PCB) levels in Lake Michigan fish. A fixed number of samples based on the variance shown in the previous 3 years of sampling should be taken and analyzed annually. Analyses of existing contaminant samples for salmonids showed that PCB levels have declined significantly during the last 10 years and are generally higher in Green Bay vs. the main lake basin. The relationship between fish length and PCB level differed significantly between years and was not found at all for some species, suggesting that length will not be a consistent predictor of PCB levels. Analysis of covariance with length as the covariate is suggested for use in establishing safe-unsafe length categories for species that show a significant length-PCB relationship over the previous 3 years of sampling. Analyses of samples from warm water species were inconclusive because these species showed very site-specific contamination levels and sample sizes from each site were then too low.

---

*Presently Systems Ecologist in the Bureau of Fish Management.

Department of Natural Resources • Madison, Wisconsin

CONTENTS

SUMMARY

This report discusses several candidate sampling designs that can be used as part of the Wisconsin Department of Natural Resources Lake Michigan fish contaminant monitoring program. Included are background information on the statistical tests involved, analyses of existing data for patterns and levels of contamination, specific sample sizes required under the various candidate methods, and recommendations on the best methods. The primary objective of this monitoring program is to accurately determine existing levels of contamination in Lake Michigan fishes. A secondary objective is to provide a data base adequate to study spatial and temporal patterns of contamination.

The current data base (1971-84) was examined for patterns and quantifiable trends. The data were generally inadequate for a statistical comparison of PCB levels at different times of the year, but PCB levels were generally higher in Green Bay vs. the main basin for the cool water fishes such as the salmonids. The PCB levels in warm water species such as carp, northern pike, and walleye were extremely variable and probably site-specific, precluding pooling over geographic areas.

Few statistically significant differences in mean PCB level between parts of fish tested were found, although an extensive analysis was not possible. In most cases, fillets had a slightly higher mean, so use of fillets should result in conservative conclusions. The PCB levels showed a significant decline in all cool water species tested, but there is substantial year-to-year variability. There was a linear relationship between length and PCB level for main basin coho, chinook, brown trout, bloater, and lake whitefish, and Green Bay lake whitefish; a logarithmic relationship for main basin lake trout; and no apparent relationship for the other species. The length-PCB relationship differed significantly between years and in some years was nonsignificant for all of these species except the brown trout and bloater. Thus length will often not be a consistent predictor of PCB content and use of length categories may lead to inefficient sampling designs. Until more information is available on annual trends, the entire sampling procedure should be repeated annually and health advisories should be based on data averaged over the preceding 3 years.

The mean contaminant level appears to be the best choice for a reporting statistic. The mean value has direct interpretation as the expected intake of contaminant if all fish are consumed. With the analysis techniques available for means, fish length can be used as a covariate and approximate safe/unsafe size categories revised as necessary without repeating the entire sample for each size category. The fixed sample design will provide a stable data base, allow for study of spatial and temporal trends, be easy to use, and allow for accurate advance allocation of resources. A fixed sample based on variances from the last 3 years should be taken annually. Recommended minimum sample sizes are given and these should be repeated for at least each lake basin.

INTRODUCTION

An issue that routinely surfaces in assessing fish contaminants is the statistical reliability of the resulting data base, particularly information on polychlorinated biphenyls (PCB) in Lake Michigan trout and salmon samples (Simmons 1984). To address this problem several candidate sampling designs that can be used as part of the Lake Michigan fish contaminant monitoring program were examined. Included are background information on the statistical tests involved, analyses of existing contaminant data (1971-84) for patterns and levels of contamination, specific sample sizes required under the various candidate methods, and recommendations on the best methods. Several candidate statistical methodologies are presented in detail to allow for changes in the program objectives or resource allocations. Although the sampling strategies could be applied to any toxics monitoring program, this study specifically addresses sport and commercial species important in Lake Michigan: coho and chinook salmon; brook, brown, rainbow, and lake trout; bloater; whitefish; carp; northern pike; walleye; white sucker, and yellow perch (scientific names for all species given in Appendix I).

OBJECTIVES OF THE MONITORING PROGRAM

The primary objective of the Lake Michigan monitoring program is to accurately determine existing levels of contamination in fishes and report those levels in a manner that allows meaningful interpretation by both the scientific and lay communities. Accuracy requires a sound statistical sampling design and a resource commitment, usually money, commensurate with the accuracy levels desired. Candidate statistics for summarizing results of the sampling include the mean contaminant level or some simple categorization scheme such as "safe" vs. "unsafe."

Information on the health risks at various levels of contamination is necessary to either evaluate the contaminant level or construct the categorization criteria. Such risk assessment is outside the scope of this report and possibly of this department, so the U. S. Food and Drug Administration standards are used. The advantage of the more detailed mean level reporting is that the public is allowed to choose their own risk levels. The categorization provides insufficient detail to evaluate risk levels beyond what is contained in the sometimes arbitrary federal standards. The reporting of the sampling should also contain auxiliary information which will assist the angler in assessing the health risk if such information is available (e.g., differences between sizes of fish, areas of the lake, or seasons in the year).

A secondary objective of the monitoring program is to provide a data base adequate to study spatial and temporal patterns of contamination. Such information can be useful not only in establishing more meaningful health advisories, but in documenting time trends, responses to mitigation measures, differences between populations, and other changes relevant to management of the contaminant problem.

CANDIDATE STATISTICAL METHODOLOGIES

The initial consideration is the statistic that will be used to report contaminant levels. The mean contaminant level is the traditional choice and has many advantages (Leidel and Busch 1975, Bar-Shalom et al. 1975). It can be directly interpreted as the actual consumption level of the contaminant if the angler eats all fish caught. The mean is easy to calculate and can be analyzed by powerful statistical techniques such as analysis of variance (ANOVA) or covariance (ANCOVA), which can directly incorporate any auxiliary variables that are found to have a quantifiable relationship with contaminant levels (Freund and Littel 1981). As will be discussed later, precise estimation of the mean will require fewer samples than precise estimation of a categorical variable. If a simple categorization is desired, the mean contaminant levels can be easily translated based on the established standards. One disadvantage is that the mean alone provides no information on the maximum contaminant levels or the actual distribution of contaminants in the population.

A categorization procedure that has been proposed is based on the proportion of individual fish in the population that exceed the established standard. This categorization has the advantage of being simple to interpret and understand; however, it has several disadvantages. The categorization relies heavily on the established standard. If the standard is revised, each fish must be recategorized and the proportion recalculated. Fish with contaminant levels near the established standard may be classed arbitrarily because of lab error or variance in the testing procedure and then contribute only random variation to the analysis. The proportions are more cumbersome to use and less accurate for evaluating the spatial and temporal changes discussed previously. The classification of the fish populations using the proportions will require a much larger sample size because simply classifying a fish as "safe" or "unsafe" throws away information on the magnitude of the difference from the standard. Since auxiliary information cannot readily be incorporated into the analyses, the entire sample will have to be repeated for each spatial, temporal, and fish size grouping.

One hybrid reporting scheme would involve a combination of the mean level and the proportion of safe fish. Such a reporting would utilize the advantages of both statistics, but would still require the maximum sample size and could lead to conflicting guidelines. For example, 11 brown trout analyzed from Lake Michigan in 1984 had a mean PCB level of 1.85 with a standard deviation of 0.7853 and a maximum observed level of 3.7. If the sample size were large enough (for a 95% confidence interval, the required sample would be 126), then the confidence interval would shrink so that the upper bound of the confidence interval would be less than 2 ppm and the population judged "safe." However, only 8 of the 11 or 72.8% were "safe" (less than 2 ppm) and regardless of how much the sample is increased this population would never be judged "safe" under the 90% safe criteria. The mean retains the advantage of having direct interpretation of the exposure level if all fish are eaten, so somebody consuming all fish caught from this population can expect to receive a dose of 1.85 parts of PCB for each million parts of fish eaten.

Regardless of the reporting statistic used, either a fixed sample or a sequential sample design can be used for the data analysis. Under both procedures the confidence levels and the precision criteria must be established by the investigator. With a fixed sample design, the number of samples to be analyzed is based on variance and/or mean estimates from preceding years' data. A fixed sample allows an accurate projection of resource use and will result in a data base with stable numbers of samples in each spatial or temporal class. Stable sample numbers will allow examination of changes and trends as discussed above. The fixed sample analysis is simpler to calculate, but on the average will require a larger sample size.

The sequential method is advantageous if the primary concern is to limit overall cost to the program and if the cost of analyzing samples is much greater than the cost of collecting field samples (Green 1979) as with the Lake Michigan fish data. Samples are collected, then analyzed one at a time or in small groups. As the results become available the statistical analyses are updated until predetermined decision criteria are satisfied, then the expensive lab analyses can be suspended. The sequential methods use fewer samples on the average than the fixed methods; however, in a small percentage of the trials the sample size will be larger (Wetherill 1966, Colton and McPherson 1976). While the sequential method offers some potential cost benefits, it is completely dependent on the specific hypotheses being tested. If the analysis is based on the established standard and the standard is revised, it is likely the current results will be insufficient to evaluate a new set of hypotheses. Further, the variable nature of the sample sizes will result in an unbalanced data base which may be insufficient to examine changes or trends in spatial or temporal factors.

Several compromises that incorporate both fixed and sequential sampling schemes may alleviate the problems associated with each. Fixing a certain minimum sample size and sequentially evaluating the predetermined hypotheses by using subsequent samples will result in a more stable data base (Billard and Vagholkar 1969; Billard 1977a, 1977b); however, the sample sizes will occasionally be lower than desired. Sequential methods could be used to evaluate specific hypotheses in most years (say 2 out of 3), while a stable data base could be built by collecting larger fixed samples in the remaining years. Species could be staggered so that the larger fixed samples do not occur in the same years. This system would be ideal if the populations do not show significant annual variations.

ANALYSIS OF EXISTING DATA

Selection of an appropriate sampling scheme will also depend on the patterns of differences between lake areas, seasons, part of fish analyzed, and years. There is some evidence that PCB concentrations differ with fish length and fat content (Simmons 1984, St. Amant et al. 1984). Thus the current data base was examined for patterns that would enable pooling of spatial or temporal factors, or quantifiable trends such as a linear relationship between PCB and fish length. The data covered the period 1971-84 and were from the ongoing Wisconsin Department of Natural Resources contaminant monitoring program as described in Pariso et al. (1984) and St. Amant et al. (1984).

Although the current data base was expensive to collect, it is variable and often contains comparatively small sample sizes. It provides limited inferential power in making contaminant safety decisions or in studying factors affecting contaminant levels. The data were taken over many variables: year, month, lake area, part of the fish analyzed, species, fish size; and there were usually too few fish for a complete statistical analysis. To permit some limited analyses, months and lake areas were pooled and analyses restricted to subsets of the data with similar time and area collections and reasonable sample sizes.

To examine spatial patterns in PCB concentrations, the lake was divided into 2 geographic regions: Lake Michigan and Green Bay. Interim analyses quickly showed that the PCB levels for some warm water species were very site specific and the two overall regions were insufficient to substantially reduce the variance. Since samples were too small to do site-specific comparisons, the remaining analyses were done only for cool water species whose migratory tendencies make the spatial subpopulations more homogeneous. Analyses of chinook were restricted to fall samples to control seasonal variability, but analyses of brown, rainbow, and lake trout, and lake whitefish were conducted on all samples because of small sample sizes. Comparisons between lake areas were further limited to years with reasonable sample sizes in both areas (Table 1).

TABLE 1. Distribution of Lake Michigan contaminant samples by species, year, and lake area.

		YEAR													
SPECIES	LAKE AREA	71	72	74	75	76	77	78	79	80	81	82	83	84	ALL
BROOK TROUT	MAIN BASIN	0	0	0	0	3	0	4	6	6	2	4	0	2	27
BROWN TROUT	GREEN BAY	0	0	0	0	0	0	0	3	0	0	0	0	12	15
	MAIN BASIN	0	0	16	12	7	0	7	9	8	6	9	6	15	95
CARP	GREEN BAY	0	0	0	7	9	5	15	7	12	19	8	9	0	91
	MAIN BASIN	0	0	0	7	0	3	18	9	7	3	6	10	5	68
CHINOOK SALMON	GREEN BAY	0	0	0	0	0	0	3	0	0	9	5	2	1	20
	MAIN BASIN	0	0	8	0	7	0	22	14	21	41	36	31	32	212
COHO SALMON	GREEN BAY	0	0	0	1	0	0	0	0	0	0	0	0	0	1
	MAIN BASIN	0	0	18	1	2	0	6	12	10	5	20	4	26	104
LAKE WHITEFISH	GREEN BAY	0	0	0	18	0	10	11	11	0	0	0	5	3	58
	MAIN BASIN	0	0	0	0	6	21	5	4	0	0	0	0	6	42
BLOATER	GREEN BAY	0	0	0	0	0	0	1	0	0	0	0	0	0	1
	MAIN BASIN	0	0	0	3	0	37	16	19	0	0	3	0	19	97
LAKE TROUT	GREEN BAY	29	0	2	26	0	0	5	3	0	0	0	2	1	68
	MAIN BASIN	0	10	28	38	27	3	29	15	7	3	10	21	22	213
NORTHERN PIKE	GREEN BAY	0	0	0	0	7	15	0	6	3	5	4	0	3	43
	MAIN BASIN	0	0	0	0	0	1	3	6	2	1	0	1	3	17
RAINBOW TROUT	GREEN BAY	0	0	0	0	0	0	0	1	0	0	1	1	4	7
	MAIN BASIN	0	0	1	1	2	1	7	11	9	3	6	0	8	49
WALLEYE	GREEN BAY	0	0	0	0	4	3	0	4	4	7	1	3	3	29
	MAIN BASIN	0	0	0	0	0	0	2	0	1	0	1	0	1	5
WHITE SUCKER	GREEN BAY	0	0	0	0	5	13	0	7	1	1	2	1	0	30
	MAIN BASIN	0	0	0	0	3	1	3	6	4	1	0	4	0	22
YELLOW PERCH	GREEN BAY	0	0	0	3	7	8	1	11	2	0	1	5	2	40
	MAIN BASIN	0	0	0	2	4	0	0	2	0	0	0	6	5	19

Mean PCB levels were compared between areas by using ANCOVA. Use of length as a covariate allowed a correction for differences in length of fish sampled and also provided a test of the similarity in the length-PCB relationship between areas (Freund and Littel 1981). The PCB levels were significantly higher in Green Bay than in the main basin of Lake Michigan in all 5 species tested (Table 2). In addition, the length-PCB relationship showed a significantly higher slope in Green Bay than in the main basin for chinook salmon and rainbow trout. These analyses include several species and a range of years, and are consistently significant, which indicates it will be advantageous to collect data from both lake areas and provide separate advisories. All further analyses were done separately for lake areas.

Over the years several different fish parts have been homogenized and sent to the lab for analysis. The fillet (skin on) is currently the only part analyzed and was the most common part analyzed in past years (Append. II). However, in some previous years, the "edible portion" and the whole fish were analyzed. Because of these differences, only years with samples from both fish parts under comparison were used to test for differences in mean PCB levels, using ANOVA. Analyses were possible for brown, lake, and rainbow trout, chinook salmon, and bloater. Most comparisons did not result in significant differences; however, the brown trout fillets had a significantly higher PCB concentration than the whole fish samples, and the bloater fillets were generally lower than either the edible portion or the whole fish (Table 3). Among the nonsignificant comparisons, the fillets had a higher mean for chinook salmon and a slightly lower mean for lake trout and rainbow trout. Note that no comparisons could be made for years after 1981. Although the differences between the sampled fish parts were variable, most analyses based on the fillets will be conservative.

If a relationship could be established between fish length and PCB levels, a more accurate categorization of individual fish would be possible. Plots of length vs. PCB showed that the smallest sampled fish of most species were generally the lowest in PCB; however, larger fish generally showed substantial variation in PCB levels. Among the cool water species, there was an approximately linear relationship between PCB levels and larger fish length for coho, chinook, lake whitefish, and possibly brown trout in the main basin (Fig. 1). The length-PCB relationship was linear on the log scale for lake trout and there was little apparent relationship with larger fish length for rainbow trout, brook trout, bloater, and brown trout from Green Bay. Most of the warm water species showed substantial variation and little apparent relationship between PCB levels and larger fish length.

Much of the variation observed in the scatterplots was found to be associated with differences in the length-PCB relationship between years. Analysis of covariance was used to test for annual differences in linear length-PCB relationships. After the analysis was restricted to years with adequate samples, the length-PCB relationship varied significantly from year to year in main basin chinook and coho salmon, main basin lake trout, and both Green Bay and main basin lake whitefish (Table 4). There was no trend in the pattern, and in some years there was no relationship at all. There were no significant differences between years for main basin brown trout or bloater, and samples were insufficient to repeat this analysis for the other species.

TABLE 2. Comparison of mean PCB levels between Green Bay and the Lake Michigan main basin for years with samples from both basins. Analysis of variance was used to test for differences and the F-statistic and associated probability level are given.

Species	Years	Mean PCB (ppm) Lake Michigan	Mean PCB (ppm) Green Bay	F-value	Prob.
Brown trout	1984	2.00	3.84	18.94	0.0001
Chinook salmon*	1981	2.94	4.96	21.38	0.0001
Lake whitefish	1976-80	1.76	4.35	12.38	0.0008
Lake trout**	1975	4.38	10.29	15.84	0.0002
Rainbow trout[a]	1984	2.05	0.55	14.22	0.0055

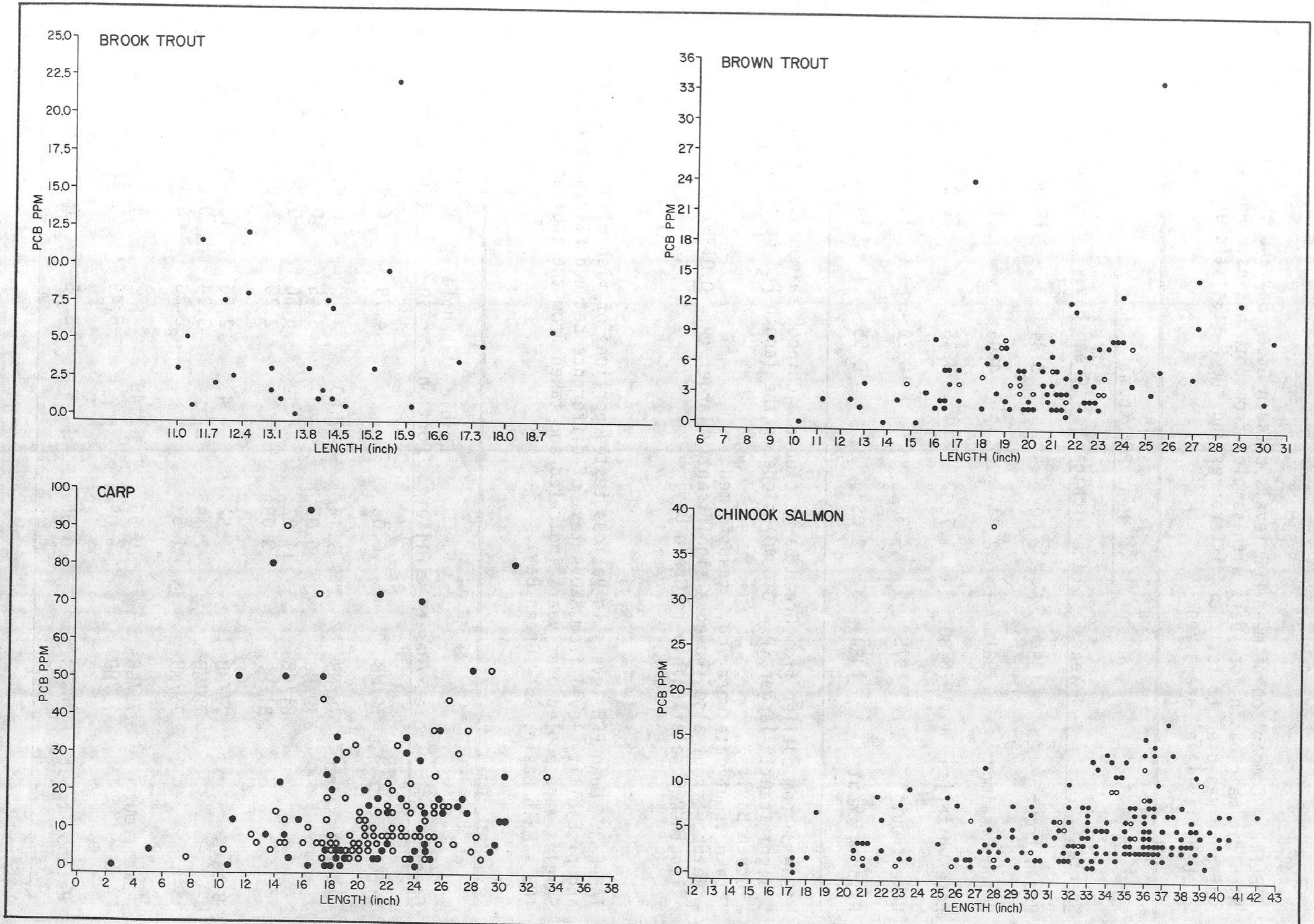
*One 38 ppm outlier from Green Bay excluded. Chinook salmon also showed a significantly larger (P = 0.0040) slope in the length-PCB relationship in Green Bay.

**Lake trout data were $\log_{10}$ transformed.

[a] Rainbow trout also showed a significantly larger (P = 0.0135) slope in the length-PCB relationship in Green Bay.

TABLE 3. Mean PCB levels in various tested portions of fish with F-statistic and probability level from ANOVA test of difference between fish portions. Samples were pooled over years indicated and taken from the indicated basin (MB = Main Basin, GB = Green Bay).

Species	Year/Basin	Fillet	Edible Portion	Whole Fish	F-value	Prob.
Brown trout	79/MB	10.03	-	4.79	6.96	0.0335
Chinook salmon	78/MB	9.16	7.60	-	0.92	0.3494
	79/MB	6.07	-	5.45	0.18	0.6824
	81/GB	11.22	-	6.56	0.34	0.5756
	81/MB	3.09	-	2.20	1.94	0.1714
Bloater	77/MB	1.52	4.43	2.06	18.48	0.0001
	78/MB	2.14	-	3.24	0.89	0.3617
	79/MB	1.63	-	2.23	8.19	0.0108
Lake trout	75/GB	10.56	12.45	-	1.18	0.2878
	75/MB	6.09	-	6.49	0.04	0.8342
Rainbow trout	78-79/MB	4.56	-	5.15	0.11	0.7403


BROOK TROUT

BROWN TROUT

CARP

CHINOOK SALMON

FIGURE 1. Plot of PCB level vs. length for Lake Michigan contaminant samples (● = Main Basin, ○ = Green Bay).

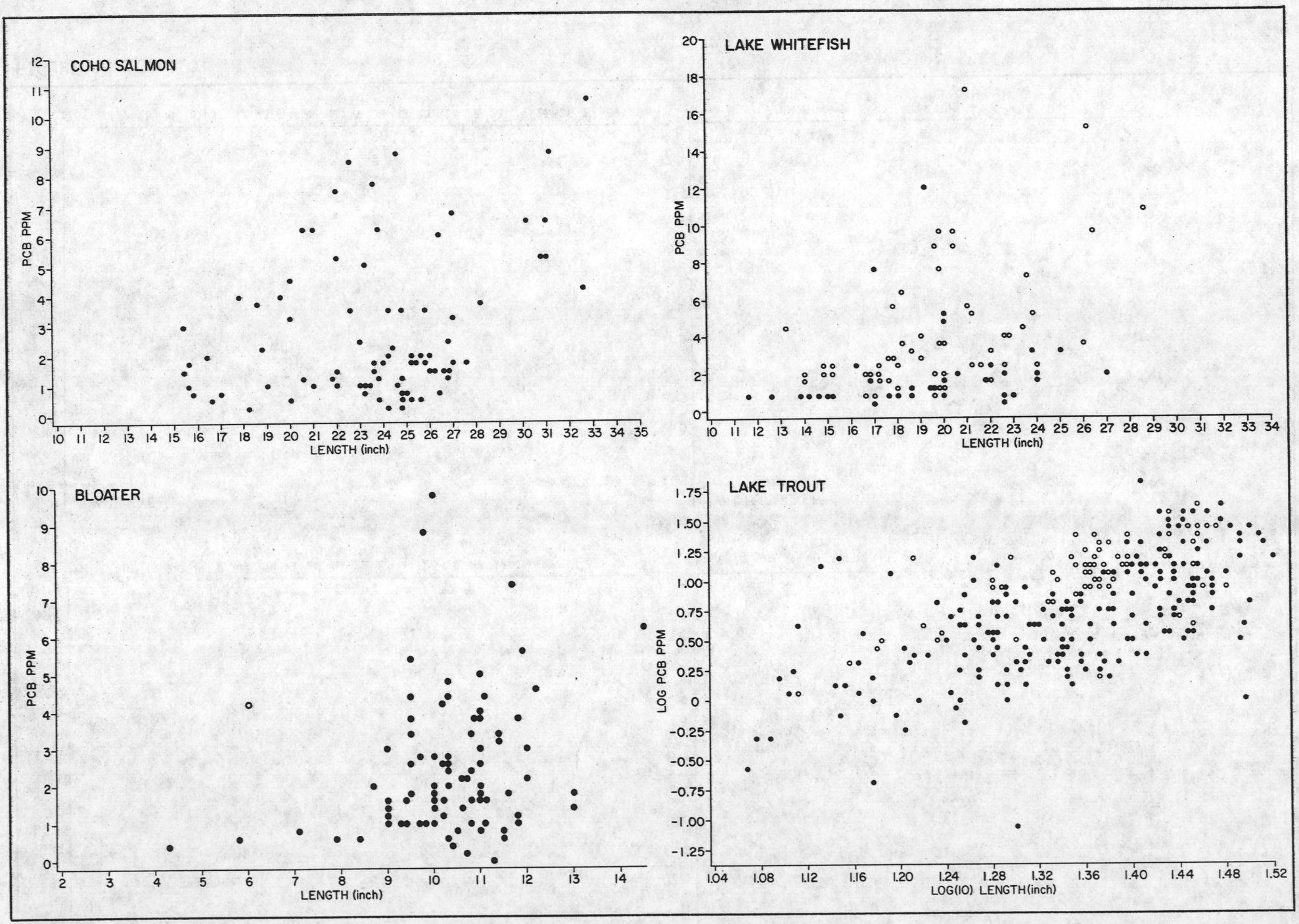


FIGURE 1. (Cont.)

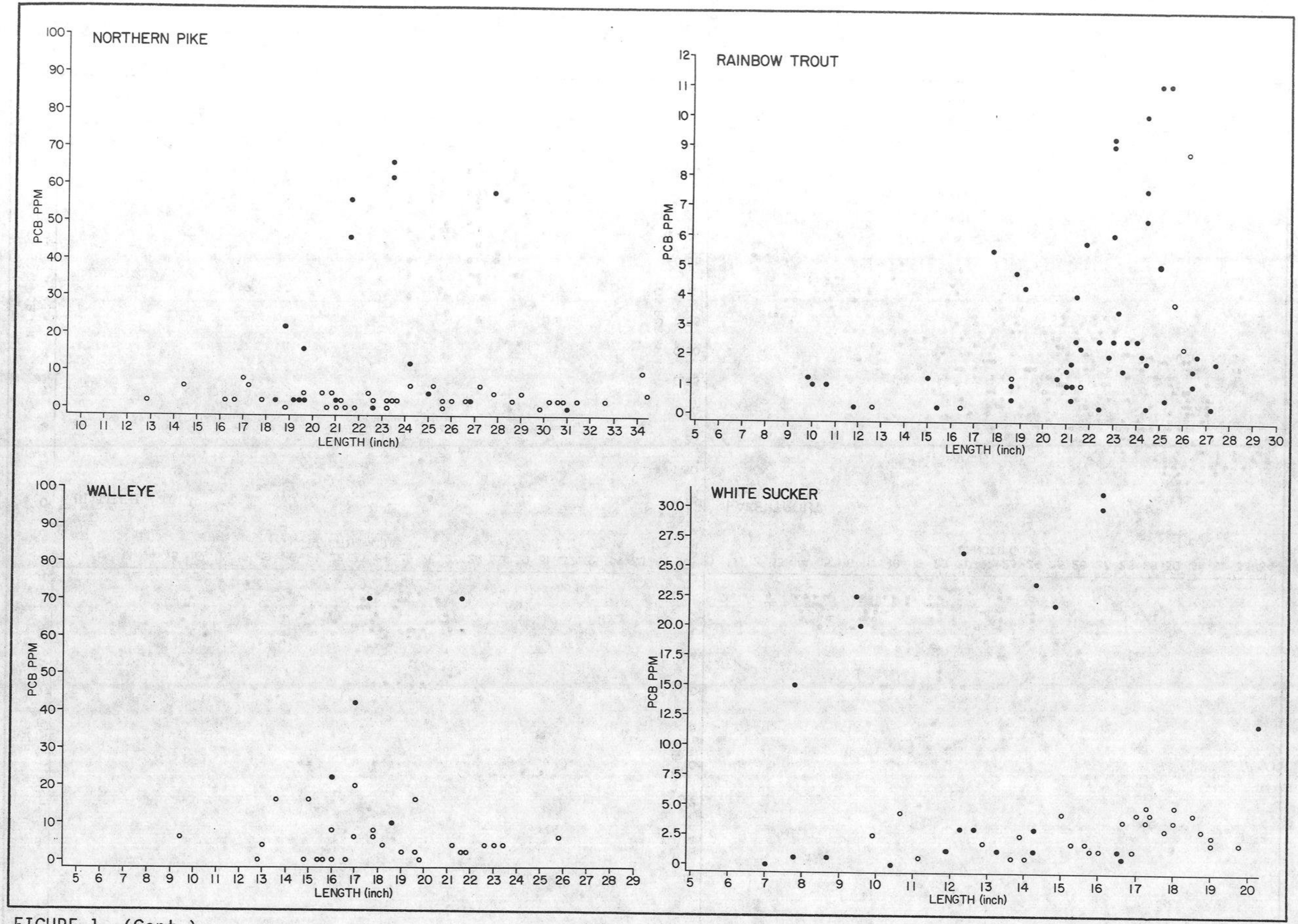


FIGURE 1. (Cont.)

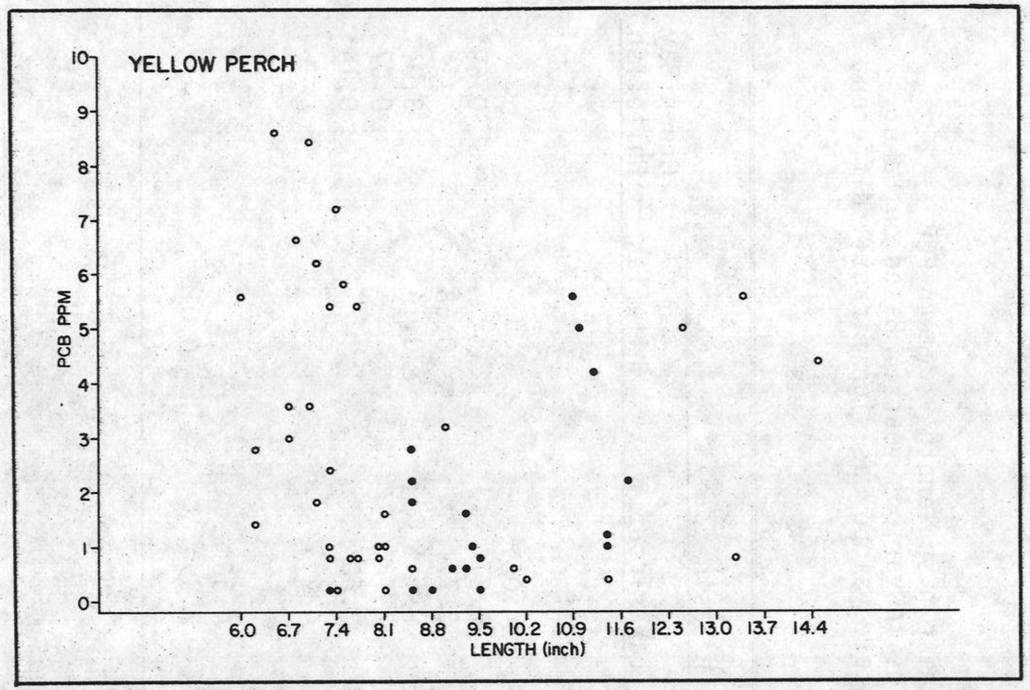


FIGURE 1. (Cont.)

These results indicate that length will not be a consistent predictor of PCB content and use of length categories, as is presently done, may lead to inefficient sampling designs and misleading conclusions. Instead, length should be entered as a covariate in an overall analysis and if a length-PCB relationship existed in any one year, the safe level cutoff length could be estimated. This design would have the advantage of not requiring separate samples for each length class.

It was of interest to see if the PCB levels have been dropping in Lake Michigan fish over the years. For species that showed different length-PCB relationships between years, predicted PCB concentrations were generated from specific years' length-PCB regression lines for several representative fish lengths. For species with a similar length-PCB relationship between years, the predicted values were generated for the average-sized fish from the pooled length-PCB regression. The predicted values were then regressed against year to test for a linear trend. There were statistically significant declines in PCB levels for all size ranges of chinook, coho, lake trout, brown trout, and bloaters, and for larger lake whitefish (Table 5). The magnitude of the decline was usually less than 1 ppm/year over the general period of 1974-84. Thus decisions relevant to current conditions should be based on current data, although the declining trend suggests that pooling data will yield conservative decisions.

TABLE 4. Comparison of slopes of linear fish length vs. mean PCB concentration (ppm/year) relationships in various years. Based on Lake Michigan samples only, asterisk indicates slope was significantly different than zero at the 5% level.

	Main Basin				
Year	Chinook	Coho	Lake Trout[a]	Whitefish	Green Bay Whitefish
1972	-	-	1.19	-	-
1973	-	-	-	-	-
1974	0.66*	0.37*	4.06*	-	-
1975	-	-	3.27*	-	0.69*
1976	0.01	-	3.77*	0.68*	-
1977	-	-	-	0.14*	3.01*
1978	0.26*	-	2.12*	-	0.07
1979	0.46*	0.01	2.60*	-	0.14*
1980	0.09	0.09	2.84*	-	-
1981	0.13*	-	-	-	-
1982	0.28*	0.16*	-	-	-
1983	0.07	-	4.56*	-	-
1984	0.08	0.06	2.99*	0.01	-

[a] Slope of $\log_{10}$ transformed relationship.

TABLE 5. Estimates of reduction in PCB levels over time. For species with significantly different length-PCB relationships size classes are separately estimated and for species with a common length-PCB relationship the estimate is for the average size fish. The estimate given is the slope of the predicted PCB levels when regressed on year and can be interpreted as ppm/year. The t-statistic and associated probability for the slope are given (MB = Main Basin, GB = Green Bay).

Species	Area	Size	Slope Estimate	t	Prob.
Chinook salmon (74-84)	MB	20"	-0.42	-3.40	0.0009
		30"	-0.77	-13.38	0.0001
		40"	-1.14	-12.33	0.0001
Coho salmon (74-84)	MB	15"	-0.21	-2.13	0.0373
		20"	-0.34	-6.22	0.0001
		25"	-0.48	-9.39	0.0001
		30"	-0.61	-6.79	0.0001
		35"	-0.75	-5.33	0.0001
Lake trout* (72-84)	MB	20"	-0.08	-10.58	0.0001
		30"	-0.06	-5.88	0.0001
		40"	-0.05	-2.88	0.0045
Lake whitefish (75-79)	GB	10"	1.55	2.98	0.0048
		15"	0.40	1.37	0.1791
		20"	-0.75	-4.44	0.0001
		25"	-1.89	-6.11	0.0001
		30"	-3.04	-5.65	0.0001
Lake whitefish (76-84)	MB	10"	0.19	1.07	0.2946
		15"	-0.11	-1.09	0.2863
		20"	-0.41	-5.30	0.0001
		25"	-0.71	-4.99	0.0001
		30"	-1.01	-4.40	0.0002
Brown trout (74-84)	MB	All	-0.13	-1.85	0.0696
Bloater (77-84)	MB	All	-0.16	-5.47	0.0001

*Log_{10} transformed data.

CURRENT CONTAMINANT LEVEL

Any statistical sampling scheme will require advance information on the contaminant level and population variance if accurate sample size estimates are needed. The PCB levels have apparently been declining in most species indicating contaminant safety decisions should be based on only the most recent data. Conversely, substantial annual variations in the length-PCB relationships and contaminant levels, and resource limitations that preclude comprehensive annual sampling, require pooling several years' data to make realistic predictions for an upcoming year. A good compromise strategy was to pool the previous 3 years' data and estimate the mean PCB levels and variance, the proportion of fish under 2 ppm PCB, and the length-PCB relationship if it existed.

Mean PCB levels were extremely variable for carp, northern pike, walleye, white suckers, and lake trout during 1982-84 (Table 6). Only brown, rainbow, and lake trout, and coho and chinook salmon had a significant length-PCB relationship during 1982-84 (Table 7). The variance of the PCB levels for lake trout was substantially reduced after accounting for differences in length; however, none of the other warm water species showed a significant length-PCB relationship. Site-specific PCB differences probably contributed substantial variance for the less migratory warm water species. For species with a significant length-PCB relationship, it was possible to predict the size of fish that would show 2 ppm PCB levels (Table 7). A different description of the contaminant level is the proportion of the population under 2 ppm during 1982-84 (Table 8). If it is assumed that the proportion of fish is binomially distributed, then the variance can be computed from the proportion as: $p(1-p)/n$, where p is the proportion and n is the number of fish examined.

TABLE 6. Mean PCB contamination levels in 1982-84 Lake Michigan samples.

			PCB (PPM)	
SPECIES	LAKE AREA	NUMBER OF SAMPLES	MEAN	VARIANCE
BROOK TROUT	MAIN BASIN	6	1.54	3.26
BROWN TROUT	GREEN BAY	12	3.84	1.87
	MAIN BASIN	30	2.86	3.60
CARP	GREEN BAY	17	16.74	345.15
	MAIN BASIN	21	12.21	328.24
CHINOOK SALMON	GREEN BAY	8	2.84	3.88
	MAIN BASIN	78	3.38	5.02
COHO SALMON	MAIN BASIN	39	1.52	1.70
LAKE WHITEFISH	GREEN BAY	8	1.83	0.42
	MAIN BASIN	6	0.90	0.14
BLOATER	MAIN BASIN	22	0.80	0.09
LAKE TROUT	GREEN BAY	3	7.33	5.69
	MAIN BASIN	53	5.62	30.47
NORTHERN PIKE	GREEN BAY	7	1.10	0.51
	MAIN BASIN	4	5.07	53.59
RAINBOW TROUT[A]	GREEN BAY	6	1.43	1.95
	MAIN BASIN	13	0.79	0.30
WALLEYE	GREEN BAY	7	6.79	61.04
WHITE SUCKER	GREEN BAY	3	1.64	0.68
	MAIN BASIN	4	6.42	32.90
YELLOW PERCH	GREEN BAY	8	0.89	0.42
	MAIN BASIN	11	0.91	0.82

[A] ONE OUTLIER WITH A PCB MEASUREMENT OF 9 PPM EXCLUDED.

TABLE 7. Estimates of population regression variance and average size of 2 ppm PCB fish (inches) for 1982-84 Lake Michigan species with a significant linear length-PCB relationship. Also given are the number of samples.

Species	Basin	Population Variance	Size of 2 ppm Fish	Number of Samples
Brown trout	Main	3.20	18.5	30
Chinook	Main	4.21	24.3	78
Coho	Main	1.34	27.0	39
Lake trout*	Main	2.62	21.4	52
Rainbow trout	Green Bay	0.67	22.7	6

*Based on $\log_{10}$ transformation.

TABLE 8. Proportion of fish under the USFDA 2 ppm standard in the 1982-84 Lake Michigan contaminant samples.

SPECIES	LAKE AREA	NUMBER OF SAMPLES	PROPORTION SAFE
BROOK TROUT	MAIN BASIN	6	0.6667
BROWN TROUT	GREEN BAY	12	0.0000
	MAIN BASIN	30	0.4333
CARP	GREEN BAY	17	0.0000
	MAIN BASIN	21	0.3333
CHINOOK SALMON	GREEN BAY	8	0.5000
	MAIN BASIN	78	0.3077
COHO SALMON	MAIN BASIN	39	0.7949
LAKE WHITEFISH	GREEN BAY	8	0.5000
	MAIN BASIN	6	1.0000
BLOATER	MAIN BASIN	22	1.0000
LAKE TROUT	GREEN BAY	3	0.0000
	MAIN BASIN	53	0.3396
NORTHERN PIKE	GREEN BAY	7	0.8571
	MAIN BASIN	4	0.5000
RAINBOW TROUT[A]	GREEN BAY	6	0.6667
	MAIN BASIN	13	1.0000
WALLEYE	GREEN BAY	7	0.4286
	MAIN BASIN	2	0.0000
WHITE SUCKER	GREEN BAY	3	0.6667
	MAIN BASIN	4	0.0000
YELLOW PERCH	GREEN BAY	8	0.8750
	MAIN BASIN	11	0.8182

[A] ONE OUTLIER WITH A PCB MEASUREMENT OF 9 PPM EXCLUDED.

SPECIFIC SAMPLING SCHEMES

FIXED SAMPLE, MEAN PCB LEVEL

Under the fixed sample scheme the formula for the necessary sample size to estimate a mean to within a desired precision is:

$$n = \frac{t^2 S^2}{d^2}$$

(Snedecor and Cochran 1978) where t is the standard normal curve ordinate corresponding to the desired confidence probability (95% = 1.96, 90% = 1.65, 80% = 1.28), S^2 is the population variance and d is the desired precision (e.g., ±0.5 ppm). The preceding formula assumes an underlying normal distribution for PCB levels. Contaminant data are often modelled with a log-normal distribution (Leidel and Busch 1975); however, the Lake Michigan PCB data appeared to be adequately approximated by a normal distribution for purposes of sample size determination.

The sample size depends on knowing S^2, the population variance, which was estimated from previous years' data (Table 6). Table 9 provides sample sizes based on the pooled 1982-84 data, and Appendix III is a more comprehensive tabulation of sample sizes necessary for various precision levels, estimated variances, and desired confidence intervals. If length is considered a covariate and the length-PCB relationship is significant, then the variances from Table 7 should be used and the sample sizes will be somewhat lower.

FIXED SAMPLE, PROPORTION SAFE

With a fixed sample design, the formula for the necessary sample size to estimate a proportion to within a desired precision is:

$$n = \frac{t^2 p(1-p)}{d^2}$$

(Cochran 1977) where t is defined as in the previous formula, p is the true proportion in the population that are safe, and d is the desired precision (e.g., ±0.001). This formula is an approximation that is not valid near proportions of 0 and 1. With very high or low proportions, tables of binomial confidence limits (e.g., Steel and Torrie 1980, Blyth and Still 1983) should be used to estimate sample sizes that will give confidence intervals of the desired width.

The sample size is proportional to the actual proportion safe, which must be estimated from previous years' data (Table 8). Table 10 contains recommended sample sizes based on the 1982-84 data, and Appendix IV presents more comprehensive tabulations of sample sizes necessary for various precision levels, estimated population proportions, and desired confidence intervals. Examination of the various estimated sample sizes makes it clear that except for the most imprecise precision levels of ±5% or ±10%, the required sample sizes to precisely estimate a proportion are impractically large. This is a key advantage of working with means. Since length cannot be readily incorporated as a covariate, the sampling must be repeated for each length grouping in addition to each spatial and seasonal class.

SEQUENTIAL SAMPLE, MEAN PCB/2 PPM STANDARD

For this design a Wald sequential probability ratio test (SPRT) is used (Wald 1947). The specific hypothesis being tested is the null hypothesis that the mean PCB level is less than 2 ppm vs. the one-sided alternative hypothesis that the mean PCB level is greater than 2 ppm. To run the test, confidence levels must be specified both for the probability of rejecting the null hypothesis when it is true (alpha error, a) and the probability of accepting the null hypothesis when it is false (beta error, b). In addition, a zone of indifference around the standard must be specified. This zone is analogous to the width of the confidence interval in a fixed sample design. For the mean sequential test, it is assumed that the population variance is known in advance, so the variance must be estimated from previous years' data.

TABLE 9. Estimated sample sizes for a fixed sample, mean PCB sampling design based on the 1982-84 uncorrected estimated variances and the variances after correcting for a significant length-PCB relationship. Confidence level was 90% and the precision is as indicated.

		1982-84	Precision (ppm)			
Species	Basin	Variance	+ 0.1	+ 0.25	+ 0.5	+ 1.0
Brook trout	Main	3.26	888	142	36	9
Brown trout	Green Bay	1.87	510	82	21	6
	Main	3.60	981	157	40	10
	Main*	3.20	872	140	35	9
Carp	Green Bay	345.15	93,968	15,035	3,759	940
	Main	328.24	89,364	14,299	3,575	894
Chinook salmon	Green Bay	3.88	1,057	170	43	11
	Main	5.02	1,367	219	55	14
	Main*	4.21	1,147	184	46	12
Coho salmon	Main	1.70	463	75	19	5
	Main*	1.34	365	59	15	4
Lake whitefish	Green Bay	0.42	115	19	5	1
	Main	0.14	39	7	2	1
Bloater	Main	0.09	25	4	1	1
Lake trout	Green Bay	5.69	1,550	248	62	16
	Main	30.47	8,296	1,328	332	83
	Main*	2.62	714	115	29	8
Northern pike	Green Bay	0.51	139	23	6	2
	Main	53.59	14,590	2,335	584	146
Rainbow trout	Green Bay	1.95	531	85	22	6
	Green Bay*	0.67	183	30	8	2
	Main	0.30	82	14	4	1
Walleye	Green Bay	61.04	16,619	2,659	665	167
White sucker	Green Bay	0.68	186	30	8	2
	Main	32.90	8,958	1,434	359	90
Yellow perch	Green Bay	0.42	115	19	5	2
	Main	0.82	224	36	9	3

*Based on variance after correcting for a significant length-PCB relationship.

TABLE 10. Estimated sample sizes for a fixed sample, percent of fish over 2 ppm sampling design based on estimated percent safe in 1982-84. Confidence level was 90% and the precision is as indicated.

		1982-84	Precision (%)				
Species	Basin	Percent	+ 0.1	+ 0.5	+ 1	+ 5	+ 20
Brook trout	Main	66.67	604,970	24,199	6,050	242	61
Brown trout	Green Bay	0*	1,000	750	250	75	30
	Main	43.33	668,513	26,741	6,686	268	67
Carp	Green Bay	0*	1,000	750	250	75	30
	Main	33.33	604,970	24,199	6,050	242	61
Chinook salmon	Green Bay	50.00	680,625	27,225	6,807	273	69
	Main	30.77	579,949	23,198	5,800	232	58
Coho salmon	Main	79.49	443,861	17,755	4,439	178	45
Lake whitefish	Green Bay	50.00	680,625	27,225	6,807	273	69
	Main	100.00*	1,000	750	250	75	30
Bloater	Main	100.00*	1,000	750	250	75	30
Lake trout	Green Bay	0*	1,000	750	250	75	30
	Main	33.96	610,581	24,424	6,106	245	62
Northern pike	Green Bay	85.71	333,451	13,339	3,335	134	34
	Main	50.00	680,625	27,225	6,807	273	69
Walleye	Green Bay	42.86	666,746	26,670	6,668	267	67
	Main	0*	1,000	750	250	75	30
White sucker	Green Bay	66.67	604,970	24,199	6,050	242	61
	Main	0*	1,000	750	250	75	30
Yellow perch	Green Bay	87.50	297,774	11,911	2,978	120	30
	Main	81.82	404,969	16,199	4,050	162	41

*Approximations based on binomial confidence interval table (1,000 was table maximum).

A good set of parameters for this contaminant scheme might be: a = 0.1, b = 0.05, and a zone of indifference of ±0.5 ppm. There is no simple formula for the required sample sizes in a sequential test, in part because the final sample size will be a random variable. Some approximate formulae have been developed to estimate the average expected sample size (Wald 1947, Sobel and Wald 1949, Colton and McPherson 1976) and these were used to estimate Table 11, which gives the expected sample sizes for the above parameters and various values of the population variance. Figure 2 shows example sequential tests for variances of 1 and 2. To use the graph in the example, the sum of the individual measurements is plotted vs. the number of samples taken. When the plot falls into one of the accept H_1 regions, the analysis is terminated. Similar sequential designs with different operating characteristics can be developed if desired. Since length cannot readily be incorporated as a covariate, the test must be repeated for each length grouping as well as each spatial and seasonal class.

SEQUENTIAL SAMPLE, PROPORTION OVER 2 PPM

The Wald SPRT is again used; however, when the hypotheses tested involve the proportion of safe fish, the underlying decision formulae must be revised. A simple test would involve a null hypothesis such as the proportion of safe fish is greater than 90% vs. the one-sided alternative hypothesis that the proportion of safe fish is less than 80%. Note that the zone of indifference is contained in the contrasting hypotheses. If a = 0.1 and b = 0.05 as before, Table 12 gives both the average sample sizes for some key proportions and the operating statistics of the test. A simple Statistical Analysis System (SAS) routine was developed to generate expected sample sizes and operating functions for a Wald proportion sequential test and can be obtained from the author. The average sample sizes are fairly small for true proportions near 0 and 1 but become relatively large for proportions in the zone of indifference.

More complex sets of hypotheses can be specified; however, the expected sample size approximations and decision functions become more complicated. Figure 3 illustrates a sequential test with the following three hypotheses: H_0: proportion safe is greater than 90%; H_1: proportion safe is between 80% and 40%; and H_2: proportion safe is 10%. The undefined areas represent zones of indifference. To use the example graph, the number of unsafe fish is plotted vs. the number of safe fish as data become available until one of the "accept" regions is reached. The expected sample numbers are shown in Figure 4 and are comparable with the previous two hypotheses test. The more complex categorization provides more information on the actual risks in fish consumption. The SAS routine described above can be used to set up more complex sequential designs also.

TABLE 11. Average expected sample sizes for a sequential design with mean = 2, a = 0.1, b = 0.05, zone of indifference = ±0.5, and the indicated assumed known variance. See Figure 2 for graph of decision functions for the variance = 1 and variance = 2 cases.*

	If true mean is:	
Known Variance	H_o: 1.5 ppm	H_1: 2.5 ppm
1.0	4.8	4.3
2.0	9.5	8.6
3.0	14.3	12.8
4.0	19.0	17.1
5.0	23.8	21.4
7.5	35.6	32.1
10.0	47.5	42.8
15.0	71.3	64.2
20.0	95.0	85.5

*Average expected sample sizes will be smaller for values less than 1.5 and greater than 2.5, and the average sample size will be larger for values between 1.5 and 2.5.

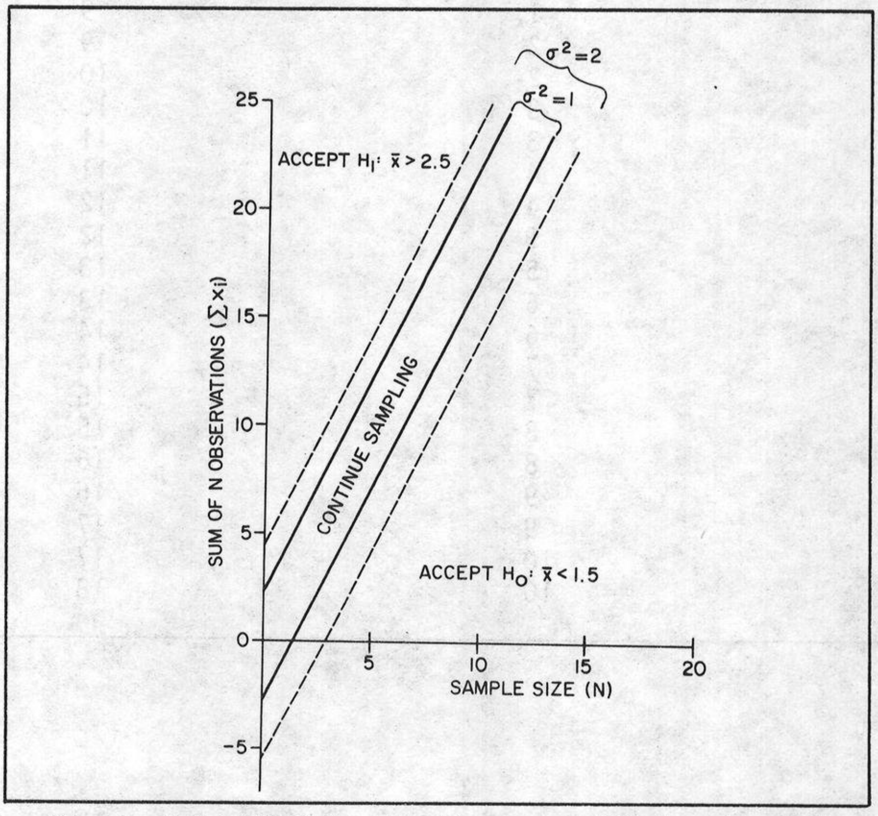


FIGURE 2. Simple Wald SPRT for testing mean PCB level vs. two alternate hypotheses: H_o: mean is less than 1.5 ppm and H_1: mean is greater than 2.5 ppm. The alpha error is 10% and the beta error is 5%. The inner bounds (solid line) are for variance = 1 and the outer bounds (dashed line) are for variance = 2.

TABLE 12. Average sample sizes and allowable number of unsafe fish for a simple Wald SPRT with H_0: proportion safe greater than 90% vs. H_1: proportion safe less than 80%. Allowable alpha error is 10% and beta error is 5%. Note that it is not possible to accept H_0 at sample sizes lower than 25.

If true P is:	Average sample size is:
100%	25
90%	65
85%	80
80%	45
0%	4

If total sample size is:	Accept H_0 if number of unsafe fish is:	Accept H_1 if number of unsafe fish is:
4	-	4
9	-	5
16	-	6
23	-	7
25	0	7
30	0	8
32	1	8
36	1	9
39	2	9
43	2	10
46	3	10
50	3	11
53	4	11
57	4	12
59	5	12
64	5	13
66	6	13
71	6	14
73	7	14
78	7	15
80	8	15
85	8	16
87	9	16
92	9	17
94	10	17
98	10	18

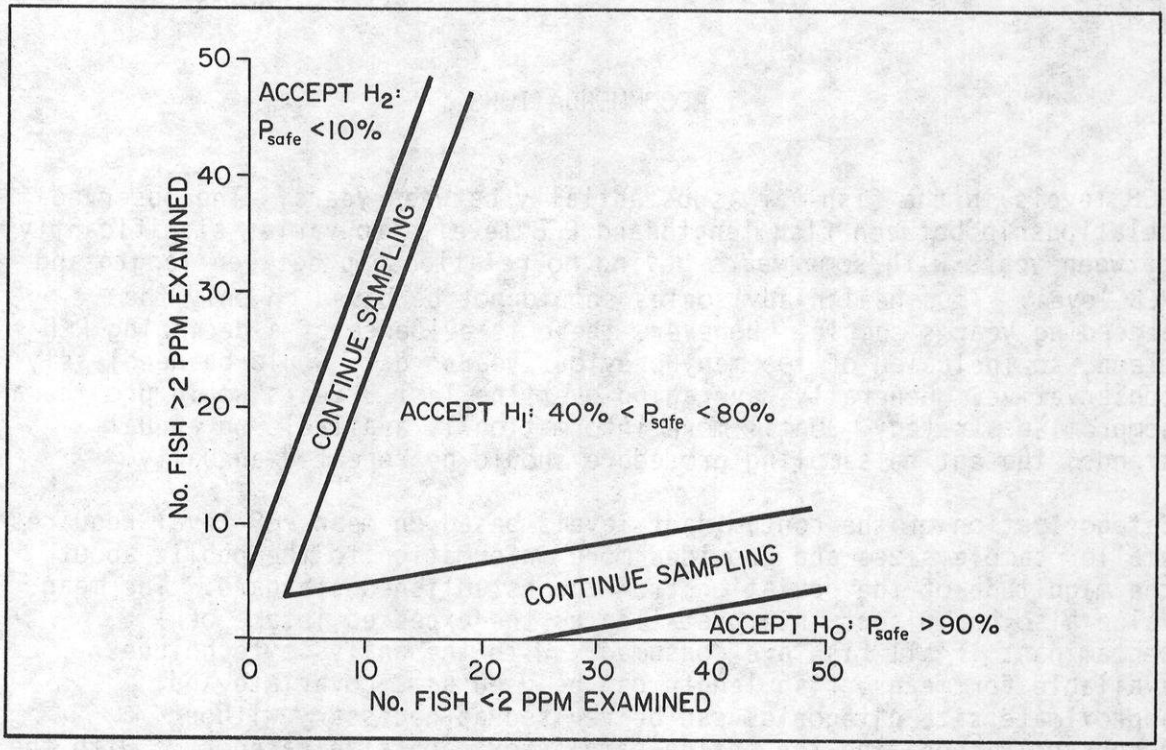


FIGURE 3. Wald SPRT for testing proportion of fish safe (P_{safe}) vs. three alternate hypotheses: H_0: $P_{safe} > 90\%$; H_1: $40\% < P_{safe} < 80\%$; and H_2: $P_{safe} < 10\%$. The alpha error rates are 10% and the beta error is 10% for H_1 vs. H_2 and 5% for H_0 vs. H_1.

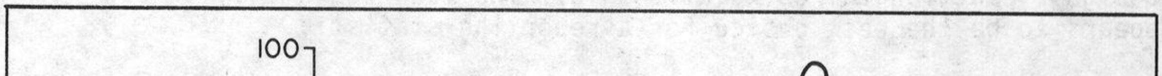


FIGURE 4. Upper bounds of average expected sample number (ASN) for the three hypotheses Wald SPRT shown in Figure 3. The ASN depends on the true underlying percent of fish under 2 ppm PCB.

RECOMMENDATIONS

PCB levels in the fish vary substantially between years. The observed relationship between fish length and PCB level also varies significantly between years with some years having no relationship between length and PCB level. Thus health advisories should not be based on only the preceding year's sample. However, there is evidence of a declining PCB trend, so inclusion of too many previous years' data would be needlessly conservative. Generally, averaging over the last 3 years would provide a compromise strategy. Until more information is available on annual trends, the entire sampling procedure should be repeated annually.

Categorization of the contaminant levels based on mean PCB level requires smaller sample sizes and provides more information to the public about the magnitude of the deviation from the established standard. The mean value also has direct interpretation as the expected intake of contaminant if all fish are consumed. With the analysis techniques available for means, fish length can be used as a covariate and approximate size categories can be revised as necessary without necessarily repeating the entire sample for each size category. With the sample size formulae given in this report and using length as a covariate, the analysis of covariance will provide the desired precision only at the average fish length in the sample. Additional samples may be necessary for increased precision at larger or smaller fish lengths. Note that additional samples can be avoided if the range of fish lengths sampled is centered on suspected safe/unsafe cutoff points or if there is no significant length-PCB relationship. The mean contaminant level appears to be the best choice for a reporting statistic.

The fixed sample design will provide a stable data base, allow for study of spatial and temporal differences and changes in contaminant levels, be easy to use, and allow for accurate advance planning and allocation of resources. If at all possible, a fixed sample based on variances from the last 3 years should be taken annually. Desired precision and confidence levels need not be overly conservative for any given year, particularly if the past several years are averaged. The recommended minimum samples based on the 1982-84 estimated variances after accounting for the different length-PCB relationship in each year are given in Table 9. The samples should be distributed approximately equally over each size range with the midrange of sampled fish centered at suspected safe/unsafe cutoff lengths. Available resources will probably dictate desired precision levels, but the 90% level and ±0.5 ppm would be conservative estimates. The significant differences found between the main basin and Green Bay for the cool water species indicate the sampling should be done separately in each area, although site-specific samples may further improve the design. Although not extensively analyzed for this report, the excessive variability noted for the warm water species suggests sampling for these species should be site-specific.

LITERATURE CITED

Bar-Shalom, Y., D. Budenaers, R. Schainker, and A. Segall
1975. Handbook of statistical tests for evaluating employee exposure to air contaminants. U.S. Dep. Health, Education and Welfare, Washington D.C., Publ. No. (NIOSH) 75-147.

Billard, L.
1977a. Optimum partial sequential tests for two-sided tests of the binomial parameter. J. Am. Stat. Assoc. 72:197-201.

1977b. Bounds for the average sample number function for a bionomial sequential test. Austral. J. Stat. 19(3):172-78.

Billard, L. and M. K. Vagholkar
1969. A sequential procedure for testing a null hypothesis against a two-sided alternative hypothesis. J. Royal Stat. Soc. (B) 32(2):285-94.

Blyth, C. R. and H. A. Still
1983. Bionomial confidence intervals. J. Am. Stat. Assoc. 78:108-16.

Cochran, W. G.
1977. Sampling techniques. Wiley and Sons, New York. 428 pp.

Colton, T. and K. McPherson
1976. Two-stage plans compared with fixed-sample and Wald SPRT plans. J. Am. Stat. Assoc. 72:80-86.

Freund, R. J. and R. C. Littel
1981. SAS for linear models. SAS series in statistical applications. SAS Inst., Inc., Cary, N.C. 231 pp.

Green, R. H.
1979. Sampling design and statistical methods for environmental biologists. Wiley and Sons, New York. 257 pp.

Leidel, N. A. and K. A. Busch
1975. Statistical methods for the determination of noncompliance with occupational health standards. U.S. Dep. Health, Education and Welfare, Washington D.C., Publ. No. (NIOSH) 75-159.

Pariso, M. E., J. R. St. Amant, and T. B. Sheffy
1984. Microcontaminants in Wisconsin's coastal zone. In J. O. Nriagu and M. S. Simmons, eds. Toxic contaminants in the Great Lakes. Wiley and Sons, New York.

Simmons, M. S.
1984. PCB contamination in the Great Lakes. In J. O. Nriagu and M. S. Simmons, eds. Toxic contaminants in the Great Lakes. Wiley and Sons, New York.

Snedecor, G. W. and W. G. Cochran
1978. Statistical methods. Iowa State Univ. Press, Ames. 593 pp.

Sobel, M. and A. Wald
1949. A sequential decision procedure for choosing one of three hypotheses concerning the unknown mean of a normal distribution. Ann. Math. Stat. 20:502-22.

St. Amant, J. R., M. E. Pariso, and T. B. Sheffy
1984. Polychlorinated biphenyls in seven species of Lake Michigan fish, 1971-1981. In J. O. Nriagu and M. S. Simmons, eds. Toxic contaminants in the Great Lakes. Wiley and Sons, New York.

Steel, R. G. D. and J. H. Torrie
1980. Principles and procedures of statistics. McGraw-Hill, New York. 633 pp.

Wald, A.
1947. Sequential analysis. Wiley and Sons, New York.

Wetherill, G. B.
1966. Sequential methods in statistics. Wiley and Sons, New York.

About the Author

During preparation of this report, Michael Staggs was Systems Analyst in the Technical Services Section of the DNR Bureau of Research. In 1986, he began a new job as Systems Ecologist with the DNR Bureau of Fish Management.

Production Credits

Ruth L. Hine and Betty L. Les, technical editors
Donna Mears, production editor
Kathryn Lyster, copy editor
Richard Burton, graphic artist
Julie McCutchin and Marlys Foley, word processors

Appendix I. Scientific names of all fish species discussed in report.

Common Name	Scientific Name
Brook trout	*Salvelinus fontinalis*
Brown trout	*Salmo trutta*
Carp	*Cyprinus carpio*
Chinook salmon	*Oncorhynchus tshawytscha*
Coho salmon	*Oncorhynchus kisutch*
Lake whitefish	*Coregonus clupeaformis*
Bloater	*Coregonus hoyi*
Lake trout	*Salvelinus namaycush*
Northern pike	*Esox lucius*
Rainbow trout	*Salmo gairdneri*
Walleye	*Stizostedion vitreum*
White sucker	*Catostomus commersoni*
Yellow perch	*Perca flavescens*

APPENDIX II. Distribution of Lake Michigan contaminant samples by species, year, lake area, and part of fish analyzed (MB = Main Basin, GB = Green Bay).

		YEAR																									
		71		72		74		75		76		77		78		79		80		81		82		83		84	
SPECIES	FISH PART	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB	GB	MB
BROOK TROUT	FILLET	0	0	0	0	0	0	0	0	0	3	0	0	0	4	0	6	0	5	0	2	0	4	0	0	0	2
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
BROWN TROUT	FILLET	0	0	0	0	0	0	0	12	0	7	0	0	0	7	0	4	0	7	0	6	0	9	0	3	12	15
	EDIBLE PART	0	0	0	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	5	0	1	0	0	0	0	0	3	0	0
CARP	FILLET	0	0	0	0	0	0	0	7	9	0	1	0	4	6	0	0	3	0	0	0	6	2	5	9	0	5
	EDIBLE PART	0	0	0	0	0	0	7	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	4	3	11	12	7	9	9	7	19	3	2	4	4	1	0	0
CHINOOK	FILLET	0	0	0	0	0	0	0	0	0	7	0	0	3	16	0	10	0	21	4	34	5	36	2	31	1	32
	EDIBLE PART	0	0	0	0	0	8	0	0	0	0	0	0	0	5	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	4	0	0	5	7	0	0	0	0	0	0
COHO	FILLET	0	0	0	0	0	0	1	1	0	2	0	0	0	5	0	11	0	10	0	2	0	20	0	4	0	26
	EDIBLE PART	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0	3	0	0	0	0	0	0
LAKE WHITEFISH	FILLET	0	0	0	0	0	0	0	0	0	6	10	21	10	5	8	4	0	0	0	0	0	0	5	0	3	6
	EDIBLE PART	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	0	0	1	0	3	0	0	0	0	0	0	0	0	0	0	0
BLOATER	FILLET	0	0	0	0	0	0	0	3	0	0	0	5	0	5	0	10	0	0	0	0	0	3	0	0	0	18
	EDIBLE PART	0	0	0	0	0	0	0	0	0	0	0	21	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	0	11	1	11	0	9	0	0	0	0	0	0	0	0	0	1
LAKE TROUT	FILLET	0	0	0	0	0	0	14	28	0	27	0	3	4	25	0	3	0	7	0	0	0	10	2	21	1	22
	EDIBLE PART	29	0	0	10	2	28	12	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	10	0	0	0	0	1	3	3	12	0	0	0	3	0	0	0	0	0	0
NORTHERN PIKE	FILLET	0	0	0	0	0	0	0	0	7	0	13	0	0	0	0	0	2	0	5	0	4	0	0	0	3	3
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	2	1	0	3	6	6	1	2	0	1	0	0	0	1	0	0
RAINBOW TROUT	FILLET	0	0	0	0	0	0	0	1	0	2	0	0	0	4	0	6	0	9	0	3	0	6	0	0	4	8
	EDIBLE PART	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	0	1	0	3	1	5	0	0	0	0	1	0	1	0	0	0
WALLEYE	FILLET	0	0	0	0	0	0	0	0	4	0	3	0	0	0	0	0	1	0	7	0	0	1	2	0	3	1
	EDIBLE PART	0	0	0	0	0	0	0	0	0	0	0	0	0	2	4	0	3	1	0	0	1	0	1	0	0	0
WHITE SUCKER	FILLET	0	0	0	0	0	0	0	0	5	3	13	0	0	0	0	0	0	0	0	0	1	0	1	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	0	1	0	3	7	6	1	4	1	1	1	0	0	4	0	0
YELLOW PERCH	FILLET	0	0	0	0	0	0	0	0	7	4	6	0	0	0	0	0	0	0	0	0	1	0	5	3	2	5
	EDIBLE PART	0	0	0	0	0	0	3	2	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	WHOLE FISH	0	0	0	0	0	0	0	0	0	0	2	0	1	0	11	2	2	0	0	0	0	0	0	3	0	0

APPENDIX III. Sample sizes necessary for estimation of a mean to within indicated precision at the indicated confidence levels.

DESIRED CONFIDENCE LEVEL=80%

ESTIMATED VARIANCE	DESIRED PRECISION (PPM)					
	0.1	0.25	0.5	1.0	2.0	5.0
0.1	17	3	1	0	0	0
0.5	83	13	3	1	0	0
1.0	166	27	7	2	0	0
1.5	250	40	10	2	1	0
2.0	333	53	13	3	1	0
2.5	416	67	17	4	1	0
3.0	499	80	20	5	1	0
3.5	582	93	23	6	1	0
4.0	666	107	27	7	2	0
4.5	749	120	30	7	2	0
5.0	832	133	33	8	2	0
6.0	998	160	40	10	2	0
7.0	1165	186	47	12	3	0
8.0	1331	213	53	13	3	1
9.0	1498	240	60	15	4	1
10.0	1664	266	67	17	4	1
20.0	3328	533	133	33	8	1
30.0	4992	799	200	50	12	2
40.0	6656	1065	266	67	17	3
50.0	8320	1331	333	83	21	3
60.0	9985	1598	399	100	25	4
70.0	11649	1864	466	116	29	5
80.0	13313	2130	533	133	33	5
90.0	14977	2396	599	150	37	6
100.0	16641	2663	666	166	42	7
200.0	33282	5325	1331	333	83	13

DESIRED CONFIDENCE LEVEL=90%

ESTIMATED VARIANCE	DESIRED PRECISION (PPM)					
	0.1	0.25	0.5	1.0	2.0	5.0
0.1	27	4	1	0	0	0
0.5	136	22	5	1	0	0
1.0	272	44	11	3	1	0
1.5	408	65	16	4	1	0
2.0	545	87	22	5	1	0
2.5	681	109	27	7	2	0
3.0	817	131	33	8	2	0
3.5	953	152	38	10	2	0
4.0	1089	174	44	11	3	0
4.5	1225	196	49	12	3	0
5.0	1361	218	54	14	3	1
6.0	1634	261	65	16	4	1
7.0	1906	305	76	19	5	1
8.0	2178	348	87	22	5	1
9.0	2450	392	98	25	6	1
10.0	2723	436	109	27	7	1
20.0	5445	871	218	54	14	2
30.0	8168	1307	327	82	20	3
40.0	10890	1742	436	109	27	4
50.0	13613	2178	544	136	34	5
60.0	16335	2614	653	163	41	7
70.0	19058	3049	762	191	48	8
80.0	21780	3485	871	218	54	9
90.0	24503	3920	980	245	61	10
100.0	27225	4356	1089	272	68	11
200.0	54450	8712	2178	544	136	22

APPENDIX III. (Cont.)

DESIRED CONFIDENCE LEVEL=95%

	DESIRED PRECISION (PPM)					
ESTIMATED VARIANCE	0.1	0.25	0.5	1.0	2.0	5.0
0.1	38	6	2	0	0	0
0.5	192	31	8	2	0	0
1.0	384	61	15	4	1	0
1.5	576	92	23	6	1	0
2.0	768	123	31	8	2	0
2.5	960	154	38	10	2	0
3.0	1152	184	46	12	3	0
3.5	1345	215	54	13	3	1
4.0	1537	246	61	15	4	1
4.5	1729	277	69	17	4	1
5.0	1921	307	77	19	5	1
6.0	2305	369	92	23	6	1
7.0	2689	430	108	27	7	1
8.0	3073	492	123	31	8	1
9.0	3457	553	138	35	9	1
10.0	3842	615	154	38	10	2
20.0	7683	1229	307	77	19	3
30.0	11525	1844	461	115	29	5
40.0	15366	2459	615	154	38	6
50.0	19208	3073	768	192	48	8
60.0	23050	3688	922	230	58	9
70.0	26891	4303	1076	269	67	11
80.0	30733	4917	1229	307	77	12
90.0	34574	5532	1383	346	86	14
100.0	38416	6147	1537	384	96	15
200.0	76832	12293	3073	768	192	31

APPENDIX IV. Sample sizes necessary for estimation of a percent to within indicated precision at the indicated confidence levels.

DESIRED CONFIDENCE LEVEL=80%

	DESIRED PRECISION (%)					
TRUE PERCENT	0.1	0.25	0.5	1.0	2.0	5.0
1	16475	2636	659	165	41	7
2	32616	5219	1305	326	82	13
3	48425	7748	1937	484	121	19
4	63901	10224	2556	639	160	26
5	79045	12647	3162	790	198	32
6	93855	15017	3754	939	235	38
7	108333	17333	4333	1083	271	43
8	122478	19596	4899	1225	306	49
9	136290	21806	5452	1363	341	55
10	149769	23963	5991	1498	374	60
11	162915	26066	6517	1629	407	65
12	175729	28117	7029	1757	439	70
13	188210	30114	7528	1882	471	75
14	200358	32057	8014	2004	501	80
15	212173	33948	8487	2122	530	85
20	266256	42601	10650	2663	666	107
25	312019	49923	12481	3120	780	125
30	349461	55914	13978	3495	874	140
35	378583	60573	15143	3786	946	151
40	399384	63901	15975	3994	998	160
45	411865	65898	16475	4119	1030	165
50	416025	66564	16641	4160	1040	166

DESIRED CONFIDENCE LEVEL=90%

	DESIRED PRECISION (%)					
TRUE PERCENT	0.1	0.25	0.5	1.0	2.0	5.0
1	26953	4312	1078	270	67	11
2	53361	8538	2134	534	133	21
3	79225	12676	3169	792	198	32
4	104544	16727	4182	1045	261	42
5	129319	20691	5173	1293	323	52
6	153549	24568	6142	1535	384	61
7	177235	28358	7089	1772	443	71
8	200376	32060	8015	2004	501	80
9	222973	35676	8919	2230	557	89
10	245025	39204	9801	2450	613	98
11	266533	42645	10661	2665	666	107
12	287496	45999	11500	2875	719	115
13	307915	49266	12317	3079	770	123
14	327789	52446	13112	3278	819	131
15	347119	55539	13885	3471	868	139
20	435600	69696	17424	4356	1089	174
25	510469	81675	20419	5105	1276	204
30	571725	91476	22869	5717	1429	229
35	619369	99099	24775	6194	1548	248
40	653400	104544	26136	6534	1633	261
45	673819	107811	26953	6738	1685	270
50	680625	108900	27225	6806	1702	272

DESIRED CONFIDENCE LEVEL=95%

	DESIRED PRECISION (%)					
TRUE PERCENT	0.1	0.25	0.5	1.0	2.0	5.0
1	38032	6085	1521	380	95	15
2	75295	12047	3012	753	188	30
3	111791	17886	4472	1118	279	45
4	147517	23603	5901	1475	369	59
5	182476	29196	7299	1825	456	73
6	216666	34667	8667	2167	542	87
7	250088	40014	10004	2501	625	100
8	282742	45239	11310	2827	707	113
9	314627	50340	12585	3146	787	126
10	345744	55319	13830	3457	864	138
11	376093	60175	15044	3761	940	150
12	405673	64908	16227	4057	1014	162
13	434485	69518	17379	4345	1086	174
14	462529	74005	18501	4625	1156	185
15	489804	78369	19592	4898	1225	196
20	614656	98345	24586	6147	1537	246
25	720300	115248	28812	7203	1801	288
30	806736	129078	32269	8067	2017	323
35	873964	139834	34959	8740	2185	350
40	921984	147517	36879	9220	2305	369
45	950796	152127	38032	9508	2377	380
50	960400	153664	38416	9604	2401	384