CALCULATION OF THE FINAL ACUTE VALUE FOR WATER QUALITY CRITERIA FOR AQUATIC ORGANISMS

by

Russell J. Erickson
Center for Lake Superior Environmental Studies
University of Wisconsin-Superior
Superior, Wisconsin 54880

Charles E. Stephan
U. S. Environmental Protection Agency
Environmental Research Laboratory-Duluth
6201 Congdon Boulevard
Duluth, Minnesota 55804

ENVIRONMENTAL RESEARCH LABORATORY-DULUTH
OFFICE OF RESEARCH AND DEVELOPMENT
U.S. ENVIRONMENTAL PROTECTION AGENCY
DULUTH, MINNESOTA 55804

DISCLAIMER

This report has been reviewed by the Environmental Research Laboratory-Duluth, U.S. Environmental Protection Agency, and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

AVAILABILITY NOTICE

This document is available to the public through the National Technical Information Service (NTIS), Springfield, Virginia 22161.

ABSTRACT

The Final Acute Value (FAV) for a material, which is an integral part of the procedure for deriving water quality criteria for aquatic organisms, is an estimate of the fifth percentile of a statistical population represented by the set of Mean Acute Values (MAV) available for the material, a MAV being the concentration of the material that causes a specified level of acute toxicity to aquatic organisms in some taxonomic group. A new procedure for calculating FAVs has been developed under the assumption that sets of MAVs are random samples of such populations. Based on examination of available sets of MAVs, it was inferred that FAV estimation would be best served by assuming that the populations have a log triangular distribution. Also, because this or any other assumption will likely not completely hold over the entire range of data in all sets, it was judged that FAV estimation should be based on a subset of the data near the fifth percentile. Based on simulations, it was determined that a FAV for a set of MAVs would be best calculated by (a) assigning each MAV a cumulative probability $P_R=R/(N+1)$ (R=rank, N=number of MAVs in the set), (b) fitting a line to ln(MAV) versus $\sqrt{P_R}$ using the four points with P_R nearest 0.05 and using the geometric mean functional relationship to estimate slope, and (c) calculating the FAV as the concentration corresponding to $P_R=0.05$ on this line. Major modifications of this new procedure were found to result either in only minor changes in FAVs or in FAVs at variance with the data. The old procedure for calculation of FAVs was judged to have some theoretical and practical shortcomings that make it less desirable than the new procedure, but FAVs by the two procedures were generally similar. A procedure based on extreme deviation from random sampling generally did not produce greatly different FAVs.

CONTENTS

TABLES

FIGURES

ACKNOWLEDGMENTS

Charles Norwood helped design this project and Robert W. Andrew performed some of the calculations.

INTRODUCTION

On November 28, 1980, the U.S. Environmental Protection Agency published "Guidelines for Deriving Water Quality Criteria for the Protection of Aquatic Life and Its Uses" as Appendix B of an announcement of the availability of water quality criteria documents (1). Calculation of the Final Acute Value (FAV) is an important part of the process described in these Guidelines. A FAV is a concentration of a material derived from an appropriate set of Mean Acute Values (MAVs), a MAV being the concentration of the material that causes a specified level of acute toxicity to an aquatic taxon in laboratory tests. The FAV is defined to be lower than all except a small fraction of the MAVs that are available for the material. The fraction was set at 0.05 (i.e., the FAV lies at the fifth percentile of the MAVs) because other fractions resulted in FAVs that were deemed too high or too low in comparison with the sets of MAVs from which they were obtained. However, if the set contains a MAV for an important species that is lower than the calculated FAV, the FAV is set equal to that MAV.

In order to be useful, the procedure for obtaining a FAV from a set of MAVs must be objective so that different parties will obtain the same FAV from a set of MAVs. The development of a reasonable mathematical framework for FAV calculation was therefore necessary. In addition, it is desirable that the rationale for the calculation procedure be relatively easy to understand and that the computations be as simple as possible. Section IV.I-O of the Guidelines described a procedure for calculating a FAV from a suitable set of Species Mean Acute Values (SMAVs). Because of criticism of this procedure, this project was initiated to define the general problem of calculating a FAV, to evaluate alternative procedures, and to recommend the most appropriate

procedure. This project was not intended to evaluate the definition of the FAV or the procedures for obtaining MAVs.

Development of an appropriate procedure for calculating FAVs requires the availability of typical sets of SMAVs. Some of the water quality criteria documents (1) contain such sets in Table 3 of the section on Aquatic Life Toxicology. Twenty data sets for freshwater species and seventeen for saltwater species were considered to be acceptable for the purposes of this project because they contained SMAVs from at least eight families in a variety of taxonomic and functional groups. These data sets (Table 1) contain from 8 to 45 SMAVs for a variety of organic and inorganic materials. Because all accepcable sets of SMAVs (that were available at the completion of this project in May, 1982) were used and because they include a diversity of species and materials, this group of 37 data sets should be representative of the data sets from which FAVs will be calculated.

There is some concern that FAVs would be more appropriately based on a taxonomic level higher than species (e.g., family). Statistical analysis of data sets similar to those in Table 1 has shown that differences between families are usually greater, often by an order of magnitude or more, than average differences within families (2). Therefore, if a set of SMAVs has a disproportionate number of species from a sensitive or insensitive family, the FAV might be undesirably affected. For example, of the 29 SMAVs for zinc in fresh water, six are from Salmonidae and are all among the twelve lowest SMAVs. Resolution of which taxonomic level is most appropriate is not of concern here, but because the definition of the FAV might be so modified, Family Mean Acute Values (FMAVs), the geometric mean of all the SMAVs available for a family, were computed for all data sets in Table 1 and are reported in Table 2. Subsequent

analysis will consider how the use of these two different taxonomic levels might affect recommendations about the procedure for calculating a FAV. This does not, however, constitute an endorsement of either species or family as the most appropriate taxonomic level.

This report will first define the problem of FAV calculation and then discuss the general methods available for estimating percentiles. Next, the example data sets in Tables 1 and 2 will be examined to determine an appropriate statistical distribution to use in the FAV calculation procedure. Simulated samples from the selected statistical distribution will then be used to determine the procedure most appropriate for calculation of the FAV. Finally, the procedure selected will be applied to the example data sets and the significance of deviations from various assumptions of the procedure will be evaluated.

TABLE 1. EXAMPLE SETS OF SPECIES MEAN ACUTE VALUES.[a]

COPPER (FRESHWATER)		DDT (FRESHWATER)		CADMIUM (SALTWATER)		CADMIUM (FRESHWATER)		TOXAPHENE (FRESHWATER)	
Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV
45	260.	42	1230.	31	50600.	29	138.	29	180.
44	150.	41	362.	30	50000.	28	135.	28	28.
43	148.	40	192.	29	21200.	27	134.	27	26.
42	145.	39	175.	28	21000.	26	133.	26	24.
41	117.	38	68.	27	19200.	25	125.	25	20.
40	91.8	37	67.	26	12200.	24	91.4	24	15.
39	47.9	36	54.	25	10100.	23	86.7	23	14.
38	46.5	35	48.	24	6600.	22	80.7	22	14.
37	35.2	34	48.	23	5290.	21	55.9	21	14.
36	23.1	33	40.	22	4100.	20	54.7	20	13.
35	22.9	32	33.	21	3940.	19	47.0	19	13.
34	21.8	31	25.	20	3800.	18	38.2	18	12.
33	20.1	30	18.	19	3500.	17	35.9	17	11.
32	18.9	29	17.	18	3440.	16	30.3	16	10.
31	14.4	28	14.	17	2930.	15	28.0	15	9.8
30	10.1	27	12.	16	2590.	14	22.3	14	9.2
29	8.41	26	10.	15	2410.	13	19.7	13	8.7
28	5.81	25	9.3	14	1800.	12	12.2	12	6.3
27	5.37	24	8.5	13	1710.	11	7.01	11	6.
26	5.00	23	8.0	12	1670.	10	3.57	10	4.2
25	4.95	22	7.8	11	1480.	9	2.87	9	4.1
24	3.97	21	7.8	10	1220.	8	1.67	8	4.
23	3.29	20	7.3	9	1080.	7	1.15	7	3.
22	2.80	19	5.0	8	760.	6	0.87	6	3.
21	2.28	18	4.9	7	645.	5	0.29	5	3.
20	2.20	17	4.3	6	320.	4	0.09	4	2.5
19	2.20	16	4.0	5	169.	3	0.04	3	2.3
18	2.13	15	3.9	4	144.	2	0.03	2	2.
17	2.13	14	3.5	3	135.	1	0.02	1	1.3
16	2.12	13	3.2	2	78.				
15	1.99	12	3.0	1	41.3				
14	1.83	11	3.0						
13	1.68	10	2.6						
12	1.42	9	2.4						
11	1.34	8	1.9						
10	1.23	7	1.9						
9	1.07	6	1.7						
8	1.02	5	1.7						
7	0.91	4	1.6						
6	0.91	3	1.4						
5	0.76	2	1.1						
4	0.55	1	0.36						
3	0.43								
2	0.28								
1	0.23								

'BLE 1. Continued

ZINC (FRESHWATER)		ENDRIN (FRESHWATER)		MERCURY (SALTWATER)		ZINC (SALTWATER)		LINDANE (FRESHWATER)	
Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV
29	2260.	28	352.	26	1680.	24	70600.	22	676.
28	1019.	27	64.	25	1260.	23	50000.	21	485.
27	732.	26	60.	24	400.	22	39000.	20	460.
26	716.	25	34.	23	315.	21	24600.	19	207.
25	708.	24	32.	22	230.	20	9460.	18	141.1
24	699.	23	5.9	21	223.	19	8100.	17	138.
23	531.	22	4.7	20	158.	18	6330.	16	90.
22	524.	21	3.1	19	116.	17	4090.	15	83.
21	413.	20	2.1	18	98.	16	3640.	14	68.
20	367.	19	1.8	17	98.	15	3380.	13	67.1
19	315.	18	1.5	16	89.	14	2440.	12	64.
18	293.	17	1.3	15	84.	13	2160.	11	55.6
17	285.	16	1.2	14	79.	12	1780.	10	48.
16	255.	15	1.1	13	70.	11	1450.	9	45.
15	172.	14	1.0	12	60.	10	1270.	8	44.
14	169.	13	0.85	11	50.	9	1000.	7	44.
13	92.8	12	0.78	10	17.	8	950.	6	40.
12	82.6	11	0.76	9	14.	7	591.	5	32.
11	81.4	10	0.75	8	14.	6	498.	4	32.
10	64.9	9	0.69	7	14.	5	400.	3	10.5
9	57.9	8	0.54	6	10.	4	321.	2	10.
8	57.6	7	0.47	5	7.6	3	310.	1	2.
7	49.3	6	0.46	4	6.6	2	290.		
6	42.0	5	0.44	3	5.6	1	166.		
5	26.2	4	0.41	2	4.8				
4	23.1	3	0.33	1	3.5				
3	21.2	2	0.32						
2	9.09	1	0.15						
1	8.89								

TABLE 1. Continued

COPPER (SALTWATER)		NICKEL (FRESHWATER)		DIELDRIN (SALTWATER)		ALDRIN (FRESHWATER)		ENDRIN (SALTWATER)	
Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV
22	600.	22	2230.	21	50.0	21	19000.	21	14.2
21	560.	21	2030.	20	34.0	20	4900.	20	12.
20	526.	20	1540.	19	31.2	19	180.	19	3.1
19	487.	19	1080.	18	23.0	18	143.	18	1.8
18	412.	18	1010.	17	19.7	17	50.	17	1.7
17	364.	17	730.	16	18.0	16	45.9	16	1.2
16	330.	16	720.	15	14.2	15	42.	15	1.1
15	181.	15	665.	14	10.8	14	34.	14	0.95
14	141.	14	659.	13	10.0	13	32.	13	0.65
13	138.	13	627.	12	8.9	12	28.	12	0.63
12	136.	12	509.	11	8.6	11	27.	11	0.6
11	129.	11	507.	10	7.0	10	27.	10	0.36
10	128.	10	457.	9	6.0	9	21.	9	0.31
9	124.	9	440.	8	5.0	8	16.	8	0.3
8	120.	8	440.	7	5.0	7	10.	7	0.3
7	86.	7	401.	6	4.5	6	9.	6	0.28
6	69.	6	388.	5	3.5	5	8.	5	0.1
5	52.	5	302.	4	2.3	4	7.4	4	0.094
4	50.	4	234.	3	1.5	3	6.1	3	0.05
3	39.	3	208.	2	0.9	2	4.5	2	0.048
2	31.	2	78.5	1	0.7	1	4.	1	0.037
1	28.	1	54.0						

HEPTACHLOR (SALTWATER)		DIELDRIN (FRESHWATER)		LINDANE (SALTWATER)		CHROMIUM(VI) (SALTWATER)		CHROMIUM(III) (FRESHWATER)	
Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV
19	194.	19	740.	19	3680.	19	105000.	18	1075.
18	198.	18	620.	18	450.	18	93000.	17	728.
17	112.	17	567.	17	103.9	17	91000.	16	633.
16	55.	16	250.	16	66.0	16	57000.	15	233.
15	50.	15	213.	15	60.0	15	32000.	14	224.
14	32.	14	130.	14	56.0	14	30500.	13	224.
13	14.5	13	41.	13	47.	13	22000.	12	224.
12	10.	12	39.	12	35.0	12	17200.	11	191.
11	8.	11	24.	11	30.6	11	15000.	10	191.
10	6.22	10	22.	10	28.0	10	10000.	9	189.
9	3.77	9	20.	9	14.0	9	7500.	8	161.
8	3.4	8	15.	8	10.0	8	6600.	7	138.
7	3.	7	10.8	7	9.0	7	6300.	6	136.
6	3.	6	8.1	6	7.3	6	4400.	5	132.
5	1.5	5	8.	5	6.28	5	4300.	4	123.
4	1.06	4	6.1	4	5.0	4	3650.	3	119.
3	0.86	3	5.0	3	5.0	3	3100.	2	47.
2	0.8	2	4.5	2	4.44	2	2000.	1	33.4
1	0.057	1	2.5	1	0.17	1	2000.		

BLE 1. Continued

HEPTACHLOR (FRESHWATER)		NICKEL (SALTWATER)		DDT (SALTWATER)		ALDRIN (SALTWATER)		CYANIDE (FRESHWATER)	
Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV
18	320.	17	350000.	17	89.	16	100.0	15	2326.
17	148.	16	320000.	16	7.9	15	36.0	14	2240.
16	101.	15	150000.	15	7.0	14	33.0	13	639.
15	81.9	14	49000.	14	6.0	13	33.0	12	431.
14	78.	13	47000.	13	4.0	12	25.0	11	318.
13	61.3	12	25000.	12	3.9	11	17.0	10	167.
12	47.3	11	17000.	11	2.0	10	13.0	9	147.
11	42.	10	9670.	10	1.8	9	12.0	8	137.
10	29.	9	7960.	9	1.6	8	9.0	7	125.
9	26.	8	6360.	8	1.1	7	8.0	6	125.
8	24.	7	2080.	7	1.0	6	7.2	5	103.
7	23.6	6	1180.	6	0.68	5	6.0	4	102.
6	13.1	5	634.	5	0.6	4	5.6	3	102.
5	7.8	4	600.	4	0.53	3	5.0	2	83.
4	2.8	3	508.	3	0.4	2	4.1	1	57.
3	1.8	2	310.	2	0.38	1	1.5		
2	1.1	1	152.	1	0.14				
1	0.9								

TOXAPHENE (SALTWATER)		CHROMIUM(VI) (FRESHWATER)		CHLORDANE (FRESHWATER)		SELENIUM (FRESHWATER)		SELENIUM (SALTWATER)	
Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV
14	1120.	14	195000.	14	190.	13	42400.	13	17348.
13	824.	13	134000.	13	82.	12	28500.	12	14651.
12	43.8	12	120000.	12	59.	11	26100.	11	9725.
11	21.	11	69000.	11	58.	10	24100.	10	7400.
10	16.	10	59900.	10	57.	9	13600.	9	4600.
9	8.2	9	59000.	9	56.	8	12600.	8	4400.
8	4.5	8	43100.	8	45.	7	10200.	7	3497.
7	4.4	7	30400.	7	40.	6	9000.	6	1740.
6	4.4	6	30000.	6	37.	5	6500.	5	1200.
5	1.4	5	25000.	5	26.	4	3870.	4	1040.
4	1.1	4	6800.	4	25.	3	1460.	3	800.
3	1.1	3	6400.	3	15.	2	710.	2	600.
2	0.5	2	3100.	2	6.3	1	340.	1	599.
1	0.11	1	67.	1	3.				

TABLE 1. Continued

ENDOSULFAN (SALTWATER)		ARSENIC (FRESHWATER)		MERCURY[b] (FRESHWATER)		SILVER (FRESHWATER)		SILVER (SALTWATER)	
Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV	Rank	SMAV
12	730.	12	41760.	11	2000.	10	5.77	10	1400.
11	157.	11	28130.	10	2000.	9	5.52	9	550.
10	7.6	10	26042.	9	2000.	8	4.11	8	500.
9	1.31	9	22040.	8	1000.	7	0.112	7	250.
8	0.83	8	18096.	7	784.	6	0.0230	6	210.
7	0.76	7	15660.	6	249.	5	0.015	5	36.
6	0.38	6	14964.	5	240.	4	0.014	4	33.
5	0.30	5	13340.	4	50.	3	0.0123	3	21.
4	0.14	4	5278.	3	20.	2	0.0121	2	20.
3	0.10	3	1348.	2	10.	1	0.00192	1	4.7
2	0.09	2	879.	1	5.				
1	0.04	1	812.						

ENDOSULFAN (FRESHWATER)		CHLORDANE (SALTWATER)	
Rank	SMAV	Rank	SMAV
10	261.	8	120.
9	88.	7	17.5
8	6.0	6	16.9
7	5.8	5	11.8
6	3.8	4	6.4
5	3.7	3	6.2
4	3.2	2	4.8
3	2.3	1	0.4
2	0.83		
1	0.34		

[a] Taken from Table 3 in the "Aquatic Life Toxicology" sections of the water quality criteria documents (1). For the purposes of this project, the Species Mean Acute Intercepts for several of the metals in fresh water were considered to be Species Mean Acute Values. All SMAVs are in µg/L.

[b] The acute value for Faxonella clypeata should have been published originally as 20 µg/L, not 0.02 µg/L (3).

TABLE 2. EXAMPLE SETS OF FAMILY MEAN ACUTE VALUES.[a]

CADMIUM (SALTWATER)		COPPER (FRESHWATER)		MERCURY (SALTWATER)		DDT (FRESHWATER)		ZINC (SALTWATER)	
Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV
25	37600.	23	260.	23	1680.	20	1230.	20	70600.
24	21200.	22	150.	22	1260.	19	92.	19	50000.
23	19200.	21	145.	21	400.	18	67.	18	39000.
22	11100.	20	117.	20	315.	17	54.	17	9460.
21	6600.	19	46.5	19	230.	16	36.	16	6330.
20	5290.	18	45.8	18	223.	15	33.	15	6330.
19	3940.	17	38.7	17	158.	14	32.	14	4090.
18	3800.	16	35.2	16	116.	13	25.	13	3640.
17	3500.	15	22.9	15	98.	12	19.	12	3380.
16	3440.	14	14.4	14	89.	11	17.5	11	2440.
15	3260.	13	10.0	13	84.	10	10.	10	2160.
14	2930.	12	3.86	12	83.	9	7.0	9	1780.
13	2410.	11	3.58	11	79.	8	4.1	8	1450.
12	1800.	10	3.56	10	60.	7	4.0	7	1000.
11	1710.	9	2.28	9	50.	6	3.2	6	543.
10	1670.	8	2.13	8	17.	5	2.4	5	525.
9	1480.	7	2.12	7	14.	4	2.3	4	400.
8	1220.	6	1.73	6	14.	3	1.7	3	321.
7	1080.	5	1.42	5	12.	2	1.6	2	310.
6	760.	4	1.34	4	6.6	1	1.3	1	166.
5	645.	3	0.99	3	6.5				
4	320.	2	0.76	2	4.8				
3	156.	1	0.30	1	3.5				
2	78.								
1	75.								

CADMIUM (FRESHWATER)		ENDRIN (FRESHWATER)		COPPER (SALTWATER)		CHROMIUM(VI) (SALTWATER)		DIELDRIN (SALTWATER)	
Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV
18	138.	17	109.	17	600.	17	105000.	16	34.0
17	133.	16	64.	16	526.	16	93000.	15	31.2
16	86.7	15	60.	15	487.	15	91000.	14	23.0
15	85.9	14	32.	14	412.	14	57000.	13	19.7
14	55.9	13	4.7	13	330.	13	32000.	12	18.0
13	54.8	12	4.3	12	268.	12	30500.	11	16.7
12	30.3	11	1.80	11	212.	11	22000.	10	14.2
11	28.5	10	1.50	10	160.	10	17200.	9	7.6
10	28.0	9	1.30	9	138.	9	15000.	8	7.0
9	19.7	8	1.0	8	136.	8	10000.	7	6.0
8	12.2	7	0.95	7	129.	7	7500.	6	5.0
7	8.86	6	0.85	6	120.	6	6600.	5	4.5
6	7.01	5	0.66	5	69.	5	6300.	4	2.8
5	2.87	4	0.65	4	66.	4	4300.	3	1.5
4	1.58	3	0.49	3	40.	3	3650.	2	0.9
3	1.15	2	0.48	2	39.	2	2970.	1	0.7
2	0.50	1	0.44	1	28.	1	2490.		
1	0.048								

TABLE 2. Continued

ENDRIN (SALTWATER)		HEPTACHLOR (SALTWATER)		LINDANE (SALTWATER)		NICKEL (FRESHWATER)		ZINC (FRESHWATER)	
Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV
16	14.2	16	194.	16	3680.	16	2230.	15	2260.
15	12.	15	188.	15	450.	15	2030.	14	1019.
14	3.1	14	112.	14	66.0	14	1540.	13	716.
13	1.7	13	55.	13	56.0	13	1080.	12	708.
12	1.1	12	21.5	12	55.9	12	730.	11	531.
11	1.1	11	10.	11	47.	11	720.	10	463.
10	0.63	10	8.	10	35.0	10	665.	9	315.
9	0.6	9	3.92	9	30.6	9	627.	8	251.
8	0.47	8	3.77	8	14.0	8	609.	7	213.
7	0.3	7	3.4	7	9.0	7	457.	6	161.
6	0.29	6	3.	6	7.3	6	446.	5	136.
5	0.1	5	3.	5	6.66	5	440.	4	92.8
4	0.094	4	1.5	4	6.28	4	401.	3	48.8
3	0.05	3	0.86	3	5.0	3	345.	2	42.0
2	0.048	2	0.8	2	5.0	2	234.	1	13.7
1	0.037	1	0.057	1	0.17	1	65.1		

ALDRIN (FRESHWATER)		DDT (SALTWATER)		NICKEL (SALTWATER)		CHROMIUM(III) (FRESHWATER)		ALDRIN (SALTWATER)	
Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV
14	9650.	14	89.	14	350000.	13	885.	13	100.0
13	180.	13	7.9	13	320000.	12	633.	12	36.0
12	143.	12	7.0	12	150000.	11	224.	11	33.0
11	50.	11	6.0	11	47000.	10	224.	10	33.0
10	27.5	10	4.0	10	35000.	9	211.	9	25.0
9	27.	9	2.0	9	17000.	8	207.	8	13.0
8	21.	8	1.6	8	9670.	7	153.	7	12.0
7	20.	7	1.4	7	7960.	6	138.	6	9.8
6	16.	6	0.87	6	6360.	5	136.	5	8.0
5	16.	5	0.68	5	2080.	4	132.	4	7.2
4	11.	4	0.6	4	1180.	3	123.	3	5.0
3	9.	3	0.53	3	600.	2	47.	2	5.0
2	8.	2	0.4	2	366.	1	33.4	1	3.7
1	7.4	1	0.14	1	310.				

ABLE 2. Continued

TOXAPHENE (SALTWATER)		DIELDRIN (FRESHWATER)		TOXAPHENE (FRESHWATER)		SELENIUM (SALTWATER)		ENDOSULFAN (SALTWATER)	
Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV
13	1120.	12	740.	12	180.	12	17348.	11	730.
12	824.	11	593.	11	28.	11	14651.	10	157.
11	43.8	10	191.	10	21.	10	9725.	9	3.16
10	16.	9	39.	9	20.	9	7400.	8	0.83
9	9.6	8	30.	8	13.	8	4600.	7	0.76
8	8.2	7	24.	7	12.0	7	4400.	6	0.38
7	4.5	6	20.	6	8.0	6	3497.	5	0.30
6	4.4	5	11.	5	5.8	5	1200.	4	0.14
5	1.4	4	8.	4	4.7	4	1180.	3	0.10
4	1.1	3	5.5	3	3.5	3	1040.	2	0.09
3	1.1	2	5.0	2	2.6	2	600.	1	0.04
2	0.5	1	4.5	1	1.3	1	599.		
1	0.11								

HEPTACHLOR (FRESHWATER)		CHROMIUM(VI) (FRESHWATER)		SELENIUM (FRESHWATER)		LINDANE (FRESHWATER)		CYANIDE (FRESHWATER)	
Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV
10	180.	10	162000.	10	42400.	10	532.	10	2326.
9	148.	9	71900.	9	28500.	9	207.	9	2240.
8	58.6	8	63800.	8	24100.	8	138.	8	431.
7	37.0	7	59900.	7	13600.	7	94.8	7	306.
6	29.5	6	30400.	6	12600.	6	68.	6	199.
5	24.8	5	30000.	5	9580.	5	53.1	5	167.
4	7.8	4	25000.	4	6500.	4	52.9	4	125.
3	2.8	3	6400.	3	6170.	3	22.4	3	118.
2	1.8	2	4600.	2	1660.	2	22.	2	83.
1	1.0	1	67.	1	340.	1	10.	1	77.

SILVER (SALTWATER)		SILVER (FRESHWATER)		MERCURY (FRESHWATER)		ENDOSULFAN (FRESHWATER)		ARSENIC (FRESHWATER)	
Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV	Rank	FMAV
10	1400.	9	5.77	9	2000.	9	261.	8	41760
9	550.	8	5.52	8	2000.	8	88.	7	29130.
8	500.	7	4.11	7	2000.	7	5.9	6	22040.
7	250.	6	0.112	6	1000.	6	3.8	5	20190.
6	210.	5	0.0230	5	784.	5	3.7	4	18096.
5	36.	4	0.015	4	244.	4	3.2	3	14130.
4	33.	3	0.013	3	32.	3	2.3	2	1794.
3	21.	2	0.0123	2	10.	2	0.83	1	879.
2	20.	1	0.00192	1	5.	1	0.34		
1	4.7								

TABLE 2. Continued

CHLORDANE (FRESHWATER) Rank	FMAV	CHLORDANE (SALTWATER) Rank	FMAV
8	190.	8	120.
7	59.	7	17.5
6	58.	6	16.9
5	44.	5	11.8
4	32.	4	6.4
3	21.	3	6.2
2	15.	2	4.8
1	6.3	1	0.4

[a] Calculated from the Species Mean Acute Values in Table 1. All FMAVs are in µg/L.

CONCLUSIONS

1. Calculation of a FAV from a typically small set of MAVs requires that the set be considered a sample from a statistical population and that the FAV be considered an estimate of the fifth percentile of that population.
2. The set of MAVs must be assumed to have been obtained from the statistical population by a specific sampling procedure; of reasonably simple sampling procedures, an assumption of random sampling appears most consistent with actual data selection,
3. Available sets of MAVs suggest that the statistical populations are highly positively skewed and that estimation would be benefitted by logarithmic transformation of MAVs.
4. Available sets of ln(MAV)s suggest that the statistical populations are significantly and variably skewed and that FAV calculation should be based on a subset of ln(MAV)s nearest the fifth percentile.
5. Available sets of ln(MAV)s suggest that FAV estimation is better served by the assumption of a triangular distribution of the populations of ln(MAV)s than by the assumption of a normal, rectangular, or biexponential distribution.
6. Simulations using a triangular distribution indicate that 'parametric' methods for percentile estimation and 'graphical' methods in which ranked data are assigned cumulative probabilities $P_R = P(E(X_R))$ produce undesired biases in the true cumulative probabilities corresponding to fifth percentile estimates.
7. Simulations also indicate that a graphical method with (a) ranked data, $X_R = \ln(MAV)$, assigned cumulative probabilities $P_R = R/(N+1)$, (b) P_R transformed to its corresponding standard variate, $Z_R = \sqrt{P_R}$, and

(c) a line fitted to Z_R versus X_R by the geometric mean functional relationship produces the least bias among alternatives examined.

8. These simulations also suggest that it is appropriate to restrict the calculation procedure to the four X_Rs with P_Rs nearest 0.05, because (a) in the absence of skewness the precision of fifth percentile estimates is little worsened by this and (b) in the presence of skewness this avoids the introduction of substantial bias.

9. The old procedure used in the 11/28/80 Guidelines has some aspects which are contraindicated either theoretically or empirically and the new procedure described here should replace it; however, FAVs calculated from example data sets by the two procedures usually do not differ by more than a factor of 2.

10. Modifications of the recommended new procedure with respect to assumed distribution, general percentile estimation method, and subset size (up to half the set size) rarely cause FAVs calculated from example data sets to vary by more than a factor of 2. Therefore, even if it is debatable whether optimal decisions were made in developing the recommended new procedure, it is unlikely that any alternative procedure, within reason, would produce substantially different results.

11. Modification of the recommended new procedure to use the entire data set often produced substantially different FAVs from example data sets, but many of these FAVs were sufficiently at variance with the lowest MAVs in the data sets to reject this modification.

12. Modification of the recommended new procedure to reflect an extreme deviation from random sampling consistently produced higher FAVs from example data sets, but the average increase was only about fifty percent;

therefore, questions about the propriety of applying methods based on random sampling to a system in which sampling is not strictly random are probably not of great importance.

13. Recommendations about FAV calculation are the same whether MAVs are for species or families.

STATEMENT OF PROBLEM

A FAV is defined as an estimate of the concentration corresponding to the fifth percentile of a suitable set of MAVs for a material; i.e., the FAV exceeds five percent of the MAVs and is exceeded by ninety-five percent. Because the number of species tested with any particular material is usually rather small, most sets of MAVs will not have a datum which can reasonably be designated as the fifth percentile; rather, the set of MAVs must be assumed to be a sample of a population (in the statistical sense) that is large enough that a fifth percentile is defined. For example, if resources permitted, the MAVs of a material for many hundreds of aquatic taxa could be determined. Such a set of MAVs could reasonably be considered to have a fifth percentile that could be obtained by inspection and is the type of statistical population of which the available sets of MAVs are assumed to be samples. Of course, the population of MAVs would need to be determined using a mix of taxa that is acceptable to toxicologists for calculating a FAV. The above assumption is inherent to the definition of the FAV and any objection to it, or modification of it, was not a subject of this project.

An additional assumption is necessary because any estimation method using a sample from a population requires that the manner in which the sample was obtained be adequately specified. The toxicity of a material is measured by many independent investigators, who select test species based on a poorly defined combination of tradition, convenience, happenstance, and intent to diversify the mix of species. All available data meeting certain quality standards (1) are incorporated into the sets of MAVs. This incorporation step does not affect the nature of the sampling process, except that a FAV will not be calculated unless the set of MAVs is of a minimum size (eight) and contains

representatives of certain categories of species (1). It is not possible to represent this process in a form suitable for applying appropriate exact estimation methods. The issue then becomes what feasible description of sampling (e.g., random, systematic) most closely approximates this process. Random sampling was selected for the following reasons:

(1) Although they meet certain minimal diversity standards, the available sets of MAVs vary markedly, and apparently haphazardly, in the species and higher taxonomic levels they contain. Such variation is not compatible with entirely systematic sampling schemes and suggests that an appropriate sampling assumption should contain a strong random element.

(2) Even where some elements of systematic sampling are evident in the available sets of MAVs, a high correlation of toxicity with these elements is usually not apparent. Without such a correlation, an assumption of systematic sampling is not particularly needed because for practical purposes it can be approximated by an assumption of random sampling.

In addition to being as much, or more, in accord with actual sampling procedures than other tractable sampling assumptions, the assumption of random sampling may be justified, in part, by noting that, in general, deviations from this assumption may occur without seriously compromising results. Methods based on random sampling do not lose all their utility if it is not possible to rigorously define a population and to formally conduct random sampling from it. The population may even be somewhat hypothetical, being defined, in part by the data selection process. Sampling may be nonrandom, but as long as the sampling process has a low enough correlation with response, results under an assumption of random sampling will not deviate by more than a certain amount from results

under more appropriate assumptions. Consideration will be given below to what errors would be introduced if random sampling were assumed for percentile estimation when sampling is actually nonrandom.

Finally, if a procedure adopted for calculating FAVs results in criteria that are somehow independently validated, the procedure can be considered to be entirely empirical and the assumptions become part of the definition of the FAV needed to produce the desired criteria. This, however, is speculative and the question remains as to whether the procedure developed here employs the most appropriate assumptions and, if not, whether this has any substantial impact on FAVs.

DESCRIPTION OF PERCENTILE ESTIMATION METHODS

Methods for estimating, from random samples, a specified percentile of a population can generally be placed into one of two categories. These categories are presented here primarily to facilitate discussion and are not meant to imply that methods in different categories do not have some important common features or are not sometimes nearly equivalent. One notable feature of any method for estimating percentiles is the need to make at least some distributional assumptions about the population from which the sample was drawn.

For methods in the first category, the parameters in the general mathematical equation for the assumed distribution are estimated from the sample by mathematical procedures formulated to produce estimates with desired properties, such as being unbiased, having minimum variance, or having maximum likelihood. The common formulas for estimation of mean and variance from a sample from a normal population is an example of such a method. Once the parameters are estimated, it is a simple matter to substitute them into the general equation for the distribution and to estimate a desired percentile. This category will be referred to as 'parametric methods'.

The second category of methods involves ranking the data in a sample and then plotting the ranked data (X_R) versus a cumulative probability (P_R) assigned to each rank (R). For calculation simplicity, plotting is usually on a coordinate system for which the cumulative form of the assumed distribution is a straight line. A line is fitted to the plotted data by eye or by some appropriate mathematical curve-fitting technique and the estimate of the desired percentile is read off the plotted line or computed from the equation for the line. This category will be referred to as 'graphical methods', although explicit graphing is never strictly necessary. In most cases, graphical methods

are not mathematically rigorous and do not produce the unbiased, maximum likelihood, or minimum variance estimates that parametric methods are designed to produce. This does not mean, however, that graphical methods will not perform adequately in practice; in fact, in some cases their performance is very similar to that of parametric methods. Furthermore, for some cases suitable parametric methods do not exist or are unreasonably cumbersome; graphical methods thus might be a very useful alternative.

For both parametric and graphical methods, discussion here will be restricted to a class of distributions which have only two parameters, these being a location parameter and a scale parameter. By this it is meant that, for each distribution type, there exists a standard distribution with standard variate denoted 'z', such that for any distribution of this type with variate denoted 'x' there exists a location parameter 'L' and a scale parameter 'S' such that $x=L+S \cdot z$. For example, for the normal distribution, the mean and standard deviation are usually used as the location and scale parameters, respectively, and the standard normal variate (also called the 'standard normal deviate' (4)) is then as usually tabulated.

Parametric Methods

The only parametric method considered here will be a general one, termed 'best linear unbiased estimation' (5,6), which can be applied to any distribution characterized by location and scale parameters. This method is 'unbiased' in that the parameter estimates will, on the average, equal the true population parameter values. It is 'linear' in that the parameter estimates are linear functions of the data. It is 'best' in that the parameter estimates have the lowest variances of all linear unbiased techniques. There may be nonlinear or biased methods that have smaller variances, but, in general, the performance,

with respect to bias and variance, of this technique cannot be much improved. Parameter estimates ($\hat{L}$, $\hat{S}$) are obtained by minimizing the value of the matrix expression:

$$(\underline{x} - \hat{L} - \hat{S}\cdot\underline{z})^T\cdot\underline{V}^{-1}\cdot(\underline{x} - \hat{L} - \hat{S}\cdot\underline{z})$$

where: N is the sample size;
$\underline{x}$ is a (Nx1) matrix consisting of the ranked sample;
$\underline{z}$ is a (Nx1) matrix of the expected values of ranked standard variates of random samples of size N from the assumed distribution;
$\underline{V}$ is the (NxN) variance/covariance matrix for ranked standard variates of random samples of size N from the assumed distribution; and
T denotes matrix transposition.

This method has the additional advantage that $\hat{x}_p=\hat{L}+\hat{S}\cdot z_p$ is also the best linear unbiased estimate for x_p, the p^{th} percentile of the population. This can be demonstrated in a variety of ways, but is most obvious when it is realized that any particular percentile could, quite legitimately, be designated the location parameter.

This method also has the advantage that it can be applied to an arbitrary subsample of the data and still produce the best linear unbiased estimate that can be obtained from that subsample. Applying the method to a subsample simply requires eliminating, from matrices $\underline{x}$, $\underline{z}$, and $\underline{V}^{-1}$, the elements referring to data not in the desired subsample. (Note: The calculation and inversion of $\underline{V}$ is not affected by these deletions; rather, deletions are made after inversion.) A notable property of such 'censoring' of data is that, if the remaining data are those nearest the percentile of interest, the variance of the percentile estimate is little worsened as the number of data used is reduced. This suggests that using all the data, other than to determine ranks and define $\underline{V}$, has relatively little utility in this kind of estimation.

The ability of this method to use only a subsample of the data has particular significance when the distribution of the population from which the

data are drawn is not perfectly characterized. For example, when concerned with the fifth percentile, deviations from the assumed distribution that are restricted to the upper part of the distribution will impact the calculations little if only the lowest few data in the sample are used. Even if the distributional assumption is violated near the fifth percentile, the impact of this violation will be reduced as the number of data formally used in the above equations is reduced, as long as the data used are those nearest the percentile of interest. Of course, the method still makes distributional assumptions about both the data used and those not used and errors will arise if these assumptions are incorrect, but as long as the distributional assumptions are not grossly violated in the range of the selected subsample, these errors will generally be minimal. The question then arises as to the optimal subsample size (n), a small size having the advantage of reducing the effects of deviation from the assumed distribution and a large size having the advantage of reducing the variance of estimates when the distributional assumptions are correct. The answer to this question is specific to the problem of concern and will be considered below.

One troubling aspect of this methodology is, ironically, its lack of bias. This is a problem because the lack of bias is in the variate rather than in the cumulative probability; i.e., in repeated sampling, $\hat{x}_p$ will average x_p, but the true cumulative probability ($P(\hat{x}_p)$) corresponding to $\hat{x}_p$ will not average p, unless cumulative probability is linearly related to variate, which is only true for rectangular distributions (simulated sampling from a variety of populations is presented below to demonstrate this point). Because the definition of the FAV is based on protecting a certain percentage of a specified taxon, this method is inappropriate; rather, a method that is unbiased with respect to the

desired cumulative probability is desired. We are aware of no published methods of this sort. It is for this reason that graphical methods are now considered.

Graphical Methods

Graphical methods for examining cumulative distributions inherently have four issues that must be resolved:

(1) Cumulative Probabilities Assigned to Ranked Data

Formulas reported (4,7,8) for calculating the cumulative probability P_R to assign to a datum X_R with rank R in a sample of size N include R/N, (R-0.5)/N, R/(N+1), and $P(E(X_R))$. Other reported formulas (8) generally are approximations of $P(E(X_R))$.

(a) P_R=R/N is not applicable here and is sometimes misapplied in the literature. This formula does not actually describe the cumulative probability to assign to a datum X_R with rank R, but rather is the cumulative probability to assign to the range X_R to X_{R+1}; thus, any specific rank is as much assigned the proportion (R-1)/N as it is R/N. Cumulative probability graphs by this method are properly a series of horizontal segments connecting the points $[X_R,R/N]$ and $[X_{R+1},R/N]$ (R=0 to N; $X_0=-\infty$, $X_{N+1}=+\infty$), usually with vertical segments connecting the points $[X_R,(R-1)/N]$ and $[X_R,R/N]$ (R=1 to N), forming a 'staircase' graph. A smooth line depicting the cumulative distribution would generally bisect the segments and pass below $[X_R,R/N]$. The error of using R/N as a point estimate for the cumulative probability assigned to a rank can also be seen by noting that it is asymmetric about the median and that, when R=N, it is indeterminate for distributions, such as the normal, whose upper limit is $+\infty$.

(b) P_R=(R-0.5)/N is apparently an attempt to select a compromise between (R-1)/N and R/N to allow a point plot of P_R versus X_R to be made. This compromise has no rigorous basis and the attempt is much better served by the two remaining formulas. Therefore, it will not be further considered here.

(c) Assigning P_R=R/(N+1) is based on this formula being the expected value of the true cumulative probability corresponding to a rank ($E(P(X_R))$=R/(N+1)), for <u>any</u> continuous distribution. It is of particular significance here because it is directed to the expected value of the cumulative probability and thus should help reduce the bias problem discussed earlier. It has additional merit in being distribution-independent. Its use will be further explored in the simulations presented below.

(d) $P_R=P(E(X_R))$ has, by definition, obvious theoretical foundations because it denotes the cumulative probability corresponding to the expected value of X_R. ($E(X_R)$ is also called 'rankit' (4)). This formula is a counterpart to $P_R=R/(N+1)$, differing by being based on the expected value of ranked data rather than the true cumulative probabilities corresponding to ranked data. Unlike $P_R=R/(N+1)$, its values are distribution-dependent. Its use will also be further explored below, but because it is based on the expected value of the variate, it is anticipeted that it will show the same problem of bias as the parametric method above.

(2) Transformation of Axes

This is dictated by the assumed distribution and by the restriction adopted here that the assumed distribution should produce a linear plot on the selected axes. In general, it is the axis against which cumulative probabilities are plotted that is transformed and the transformation is based on the standard distribution of the assumed distribution; in fact, this can be treated as a transform of P_R to a corresponding standard variate Z_R. In such a case, the plot becomes one of a Z_R assigned to each rank versus the observed datum X_R. The slope dX/dZ is the scale parameter and the intercept on the X axis is the location parameter.

(3) Subsample Size

This issue is identical to that discussed for the parametric method. A later section will consider how the subsample size (n) can best be determined, based on simulations under various assumptions.

(4) Fitting a Line to Plotted Data

Because the restriction of a linear plot has already been made, this issue reduces to how to compute the slope of the line most appropriate to the data. Because, for any N, the Z_R or P_R assigned to a ranked datum is fixed, and thus may be an analogy to an independent variable, and because the line to be fitted can be expressed as $X_R=L+S\cdot Z_R$, it may be thought that the standard least-squares regression formula with X_R as the dependent variable and Z_R as the independent variable would be the preferred choice. As will be seen below, this turns out to be the case when $P_R=P(E(X_R))$ is used and when $\hat{Y}_p$, rather than $P(\hat{Y}_p)$, is desired to be unbiased. As before, achieving an unbiased $P(\hat{Y}_p)$ is not amenable to exact techniques and an empirical approach must be used. To this end, three different, but simple, slope formulas were considered (again, this approach is strictly empirical, employing these formulas as representing a range within which a reasonable slope might lie; nothing is implied here about a theoretical justification for one or the other formula and it is not implied that this application meets the assumptions on which any of the formulas are based):

(a) LS-X - standard bivariate least-squares with X_R as the dependent variable (residuals minimized in X-direction):

$$\hat{S} = \frac{\sum^{n} Z_R X_R - (\sum^{n} Z_R)(\sum^{n} X_R)/n}{\sum^{n} Z_R^2 - (\sum^{n} Z_R)^2/n}$$

(b) LS-Z - standard bivariate least-squares with Z_R as the dependent variable (residuals minimized in Z-direction):

$$\hat{S} = \frac{\sum^{n} X_R^2 - (\sum^{n} X_R)^2/n}{\sum^{n} Z_R X_R - (\sum^{n} Z_R)(\sum^{n} X_R)/n}$$

(c) GMFR - geometric mean functional relationship (residuals are minimized in the direction of the arithmetic reciprocal of the slope; this method produces the geometric mean of the slopes by the two previous methods and has seen some application in regression where both variables are in error (9,10)):

$$\hat{S} = \sqrt{\frac{\sum^{n} X_R^2 - (\sum^{n} X_R)^2/n}{\sum^{n} Z_R^2 - (\sum^{n} Z_R)^2/n}}$$

Whatever slope formula is used, the line always passes through the mean X_R and the mean Z_R. The location parameter estimate $\hat{L}$, which is the intercept on the X axis, is therefore

$$\hat{L} = (\hat{S} \cdot \sum^{n} Z_R - \sum^{n} X_R)/n.$$

SELECTION OF DISTRIBUTION

All methods for estimating the fifth percentile of a population from a sample require at least some assumptions about the distributional characteristics of the population. Few data sets from which an FAV will be calculated will be large enough that such characteristics can be inferred from the individual data set. However, the large number of sets available (Tables 1 and 2) provides an opportunity for evaluating these characteristics and for determining which characteristics can be reasonably applied to all data sets and which parameters must be estimated individually from each set.

It is desirable to keep the number of unknown distributional parameters as low as possible, not only because analysis becomes markedly more complicated as the number of parameters increases, but also because data sets of the minimum size (N=8) may be overly fitted if the number of parameters is not small. The example data sets (Tables 1 and 2) vary widely in their means and coefficients of variation. Therefore, at least two parameters, a location parameter (e.g., mean) and a scale parameter (e.g., standard deviation), are required.

Because these two parameters relate to the first and second moments of the samples, an obvious third parameter to consider is skewness, which is related to the third moment of the samples. Skewness is also strongly indicated by inspection of the example data sets. A skewness measure (4), the normalized third central moment, was estimated for each example data set. All sets showed positive skewness. The skewness was substantial enough to reject, at the 0.10 level of significance, the hypothesis that the set was a random sample from a normally distributed population for 35 of 37 SMAV sets and for 34 of 37 FMAV sets; at the 0.01 level of significance, this hypothesis was rejected for 30 SMAV sets and 25 FMAV sets.

Because of this strong positive skewness, a logarithmic (base e) transformation was applied to each MAV, so that discussion will now relate to the distribution of ln(MAV). The skewness measure for each data set was recomputed and the average skewness decreased from 2.39 for SMAVs and 2.08 for FMAVs to 0.06 for SMAVs and 0.07 for FMAVs.

The small average skewness does not, however, mean that individual sets can be considered to be samples from nonskewed populations. When the skewness measures of individual data sets were tested under the same null hypothesis as above, the hypothesis was rejected at the 0.10 significance level for 8 SMAV sets and 7 FMAV sets and at the 0.01 signficance level for 3 of the SMAV sets and 2 of the FMAV sets. Although this is substantially fewer than before logarithmic transformation, it still indicates that skewness in some sets might be too large to ignore. Furthermore, among the sets with significant skewness, the skewness was sometimes positive and sometimes negative, indicating that the populations these sets represent vary substantially in skewness.

Therefore, despite logarithmic transformation, skewness in the data must still be dealt with by the methodology adopted for the estimation of the fifth percentile. Two general approaches were considered for this. First, distributions with a third parameter that affects skewness and which can be estimated from a sample could be used. This approach greatly increases the difficulty of parameter estimation and it is questionable whether the smaller data sets reliably have enough information to make this effort appropriate or worthwhile. The second approach is to limit the estimation method to a subset of the data near the percentile of interest. By doing this, the effects of having non-zero skewness are markedly reduced and the location and scale parameter estimates apply only locally, incorporating the effects of skewness at

that locality. This approach allows the use of relatively simple estimation methods and, as will be further discussed below, has very little impact on the precision of fifth percentile estimates even if a population is not skewed. The second approach will therefore be employed here.

Higher moments of the data sets were not directly examined because (a) the decision to limit analysis to a subset of the data makes such an examination complicated and (b) the effects of higher moments should be adequately accounted for either by this limitation or by the examination of specific distributions that follows.

Inference of distributional characteristics from the example data sets was therefore limited to symmetric distributions with just location and scale parameters to be estimated; furthermore, the most relevant information in the data sets is that nearest the fifth percentile. The approach followed here was to examine the fit of specific distributions to the example data sets. Four distributions were considered:

(1) Rectangular Distribution

This was included as an extreme case because it assumes that the relative frequency of ln(MAV)s remains constant between some lower and upper limits, whereas theoretical considerations and inspection of the data sets suggest that the frequency declines as the lower and upper limits are approached (i.e., very sensitive and very resistant taxa are rarer than those with moderate sensitivity). The standard probability density function for this distribution is:

$$f(z) = 1/\sqrt{12} \quad ; \quad -\sqrt{3} < z < \sqrt{3}$$
$$f(z) = 0 \quad ; \quad z < -\sqrt{3}, \; z > \sqrt{3}$$

(2) Triangular Distribution

This was included because it is the simplest distribution that incorporates two basic properties that the frequency of ln(MAV)s should have: (a) sensitivity should have lower and upper limits (no species succumbs to infinitesimal concentrations of a material and no species tolerates infinite concentrations) and (b) the frequency of ln(MAV)s should decline to zero as

the limits are approached (fewer species are near the limits than are near the midrange). The standard probability density function for this distribution is:

$$f(z) = (1-|z|)/\sqrt{6} \quad ; \quad -\sqrt{6} < z < \sqrt{6}$$

$$f(z) = 0 \quad ; \quad z < -\sqrt{6}\ ,\ z > \sqrt{6}$$

(3) Normal Distribution

This was included due to its broad applicability and to provide a curved alternative to the linear frequency trend of the triangular distribution; this curvature causes relatively rare sensitive or resistant taxa to have somewhat more extreme ln(MAV)s (relative to the range of the majority of the taxa with moderate sensitivity) than does the triangular distribution. No lower or upper limits exist, but the frequency becomes so small at reasonably moderate deviations from the mean that this deficiency is probably of limited consequence. The standard probability density function for this distribution is:

$$f(z) = \frac{1}{\sqrt{2\pi}}\, e^{-z^2/2} \quad ; \quad -\infty < z < +\infty$$

(4) Biexponential Distribution

This was included as an extreme case in which the most sensitive and resistant taxa have greatly different ln(MAV)s than the majority of the taxa with moderate sensitivity. The standard probability density function for this distribution is:

$$f(z) = \frac{1}{\sqrt{2}}\, e^{-\sqrt{2}|z|} \quad ; \quad -\infty < z < +\infty$$

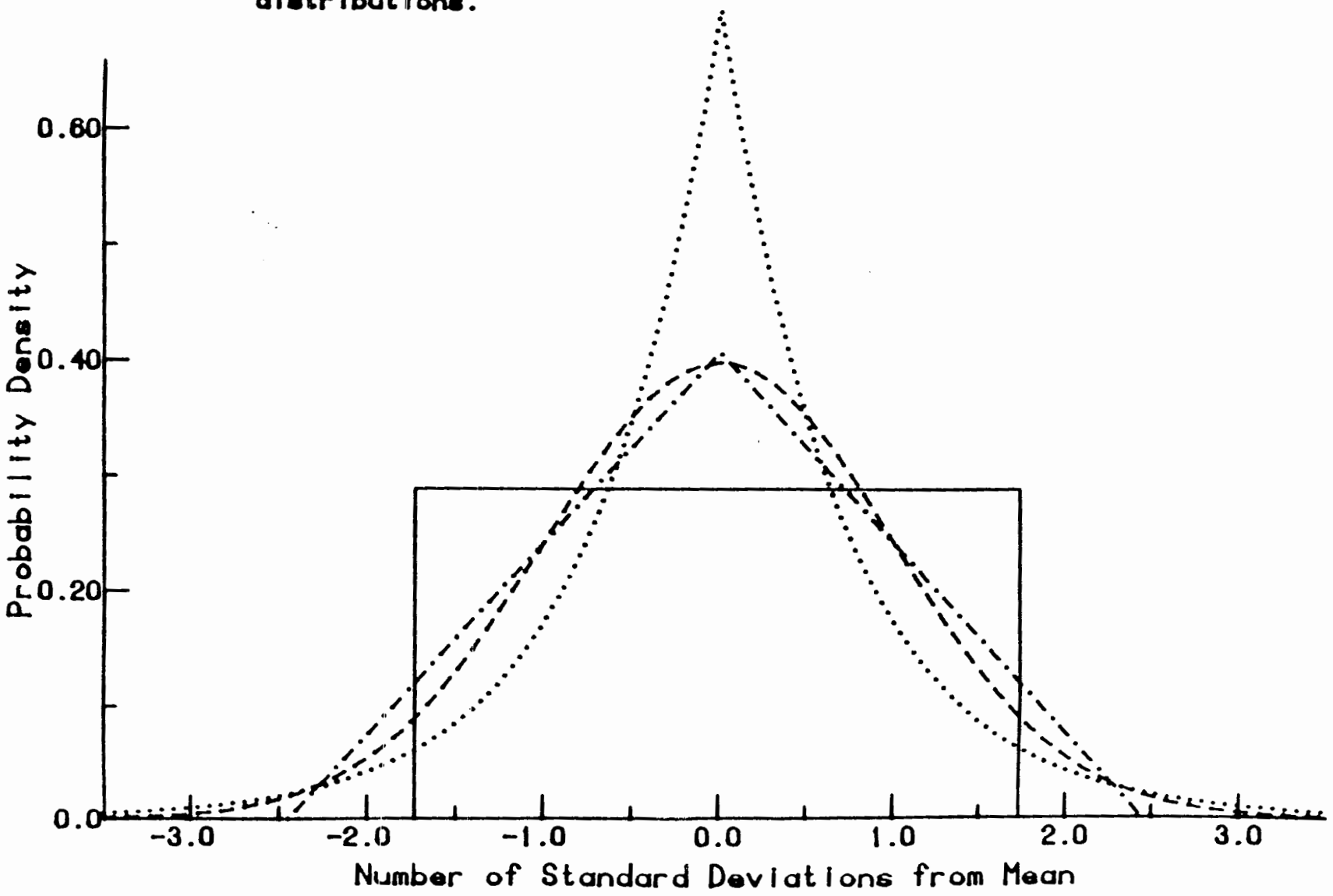
Shapes of these distributions when they have mean = 0 and standard deviation = 1 are displayed on Figure 1.

The best linear unbiased estimation method[1] discussed earlier was used to estimate location and scale parameters from each example data set for each combination of the four distributions above and four subset sizes (n=4, N/4, N/2, and N, where N=data set size; n also was required to be at least 4, which

---

[1] The matrix $\underline{V}$ for this method was calculated by exact integrals for all distributions except normal, for which approximate formulas were used (11).

Figure 1. Standard probability density plots for rectangular (———), triangular (·–·–·), normal (– – – –), and biexponential (·····) distributions.



was considered a minimum number to use to test distribution fits). From each such estimation, the expected value ($E(X_R)$) of each ranked datum was estimated as $\hat{L}+\hat{S}\cdot E(Z_R)$, where $\hat{L}$ is the location parameter estimate, $\hat{S}$ is the scale parameter estimate, and $E(Z_R)$ is the expected value of the datum of rank R in samples of size N from the assumed standard distribution. The ratio:

$$\frac{\sum^{n}(X_R-E(X_R))^2}{\sum^{n}(X_R-\bar{X}_R)^2}$$

(i.e., the fraction of the variance of the subset not explained by fitting the data to the distribution) was adopted as a measure of goodness-of-fit of the data to the assumed distribution over the size (n) of the subset used. Average goodness-of-fits for all SMAV sets are reported in Table 3 and for FMAV sets in Table 4.

The triangular distribution was selected for use in further development of the FAV calculation procedure based on its superior average goodness-of-fit to the data, especially for the subsets restricted to be near the fifth percentile. This distribution has the additional advantage of most simply embodying the two distributional characteristics that are theoretically and empirically most sensible (i.e., the existence of finite limits and a probability density function that declines to those limits), and thus constitutes a reasonable null hypothesis that should be used unless clearly rejected. Also, it constitutes a compromise distribution with a shape that is intermediate in the range exhibited by the example data sets, thus limiting potential errors. The triangular distribution has the further advantage of having simple mathematical formulations for the percentile estimation methods discussed above.

TABLE 3. AVERAGE GOODNESS-OF-FITS[a] FOR ln(SMAV) DATA SETS.

ASSUMED DISTRIBUTION	SUBSET SIZE (n) N	N/2	N/4[b]	4
RECTANGULAR	0.169	0.150	0.144	0.135
TRIANGULAR	0.082	0.099	0.114	0.118
NORMAL	0.081	0.104	0.134	0.137
BIEXPONENTIAL	0.106	0.230	0.235	0.206

[a] "Goodness-of-fit" is the fraction of variance of 'n' data points not explained by fitting data to assumed distribution; lower values indicate better fits; values should be compared only within columns.

[b] $n = 4$ when $N \leq 16$.

TABLE 4. AVERAGE GOODNESS-OF-FITS[a] FOR ln(FMAV) DATA SETS.

ASSUMED DISTRIBUTION	SUBSET SIZE (n) N	N/2	N/4[b]	4
RECTANGULAR	0.145	0.127	0.121	0.125
TRIANGULAR	0.095	0.103	0.108	0.116
NORMAL	0.087	0.120	0.129	0.143
BIEXPONENTIAL	0.109	0.277	0.218	0.238

[a] "Goodness-of-fit" is the fraction of variance of 'n' data points not explained by fitting data to assumed distribution; lower values indicate better fits; values should be compared only within columns.

[b] $n = 4$ when $N \leq 16$.

SELECTION OF PERCENTILE ESTIMATION METHOD AND SUBSET SIZE

Ten thousand computer-generated random samples from a standard triangular distribution (location parameter = mean = 0; scale parameter = standard deviation = 1) were used to estimate the fifth percentile of the distribution for each combination of the following methods, sample sizes, and subsample sizes.

(1) Percentile Estimation Methods

Seven methods were examined. These included one parametric method (best linear unbiased estimation) and six graphical methods (all possible combinations of the two formulas for assigning cumulative probabilities ($P(E(X_R))$, $E(P(X_R))$ and the three formulas for computing slope (LS-X, LS-Z, GMFR)).

(2) Sample Sizes

Sample sizes (N) of 8, 15, and 30 were selected as being representative of the minimum size, a moderate size, and a large size that are found in the available sets of SMAVs and FMAVs.

(3) Subsample Sizes

Subsample sizes (n) of the 4, N/4, N/2, and N points closest to the fifth percentile were considered; for N/4, an additional restriction of $n \geq 4$ was imposed; $n=4$ was considered to be the minimum reasonable size, a lesser number making analysis too sensitive to a spurious datum.

From location and scale parameter estimates, the estimate of the fifth percentile was calculated as $\hat{x}_5 = \hat{L} - 1.675 \cdot \hat{S}$, -1.675 being the fifth percentile for the standard triangular distribution. The average $\hat{x}_5$ over the 10,000 simulations was designated as $\bar{x}_5$ and should equal -1.675 for methods unbiased with respect to the variate. Because the parameters of the population from which the samples were drawn are known, the true cumulative probability $P(\hat{x}_5)$ corresponding to each $\hat{x}_5$ was calculated. The average $P(\hat{x}_5)$ over the 10,000 simulations was designated $\bar{P}_5$ and should equal 0.050 for methods unbiased with respect to cumulative probability. $\bar{x}_5$ is tabulated in Table 5 and $\bar{P}_5$ is

tabulated in Table 6. Table 5 also includes the standard deviations for $\hat{x}_5$ in order to indicate the relative precision of the various methods.

As expected, the best linear unbiased estimation method did produce an essentially unbiased $\overline{x}_5$, as did the graphical method using $P(E(X_R))$ to assign cumulative probability and LS-X to calculate slope (Table 5). However, it is bias in $\overline{P}_5$ that is of paramount concern here. The best linear unbiased estimation method and all graphical methods using $P(E(X_R))$ were substantially more biased than the graphical methods using $E(P(X_R))$ (Table 6) and were therefore dropped from consideration. In addition, the standard deviations of $\hat{x}_5$ by the best linear unbiased method were usually no better than 10% less than those of the graphical methods using $E(P(X_R))$ (Table 5), indicating that the better precision of the best linear unbiased method is of little consequence.

Although they did have lower biases than the other methods, none of the graphical methods using $E(P(X_R))$ to assign cumulative probability had an unbiased $\overline{P}_5$ and the bias varied with n and N (Table 6). Furthermore, none of the formulas for calculating slope had the lowest bias for all combinations of n and N. The geometric mean functional relationship was selected as having the lowest average bias over all combinations.

Selection of the most appropriate subsample size required consideration of the precision of $\hat{x}_5$ (Table 5) for the selected percentile estimation method (graphical method using $E(P(X_R))$ to assign cumulative probability and GMFR to calculate slope). For N=8, the standard deviation of $\hat{x}_5$ (Table 5) for n=4 (=N/4, =N/2) was only 12% greater than that for n=N. For N=15, the standard deviation of $\hat{x}_5$ for n=4 (=N/4) was only 1% greater than that for n=N/2 and only 6% greater than that for n=N. For N=30, the standard deviation of $\hat{x}_5$ was

TABLE 5. MEANS AND STANDARD DEVIATIONS[a] OF ESTIMATES OF FIFTH PERCENTILE ($\hat{x}_5$) BY VARIOUS METHODS, FOR 10,000 SAMPLES FROM A STANDARD TRIANGULAR DISTRIBUTION.

N	n	METHOD						
		PARAMETRIC	GRAPHICAL					
			$P_R = P(E(X_R))$			$P_R = E(P(X_R))$		
			LS-X	LS-Z	GMFR	LS-X	LS-Z	GMFR
8	4	-1.68 (0.57)	-1.68 (0.57)	-1.79 (0.61)	-1.73 (0.59)	-1.83 (0.62)	-1.95 (0.67)	-1.89 (0.64)
	8	-1.68 (0.52)	-1.68 (0.53)	-1.81 (0.55)	-1.74 (0.54)	-1.83 (0.56)	-1.98 (0.58)	-1.90 (0.57)
15	4	-1.67 (0.39)	-1.67 (0.39)	-1.73 (0.40)	-1.70 (0.39)	-1.77 (0.41)	-1.84 (0.44)	-1.80 (0.42)
	8	-1.67 (0.38)	-1.67 (0.39)	-1.75 (0.40)	-1.71 (0.39)	-1.77 (0.41)	-1.85 (0.43)	-1.81 (0.42)
	15	-1.67 (0.36)	-1.67 (0.38)	-1.76 (0.39)	-1.71 (0.38)	-1.77 (0.39)	-1.86 (0.40)	-1.81 (0.39)
30	4	-1.67 (0.25)	-1.67 (0.25)	-1.69 (0.25)	-1.68 (0.25)	-1.73 (0.26)	-1.75 (0.26)	-1.74 (0.26)
	8	-1.67 (0.25)	-1.67 (0.26)	-1.71 (0.26)	-1.69 (0.26)	-1.73 (0.26)	-1.77 (0.27)	-1.75 (0.27)
	15	-1.67 (0.25)	-1.67 (0.26)	-1.72 (0.26)	-1.70 (0.26)	-1.73 (0.27)	-1.77 (0.27)	-1.75 (0.27)
	30	-1.67 (0.24)	-1.67 (0.26)	-1.72 (0.26)	-1.70 (0.26)	-1.73 (0.27)	-1.77 (0.27)	-1.75 (0.27)

[a] Standard deviations of estimates in parentheses.

TABLE 6. MEAN TRUE CUMULATIVE PROBABILITIES ($\bar{P}_5$) OF ESTIMATES OF FIFTH PERCENTILE ($P(\hat{x}_5)$) BY VARIOUS METHODS, FOR 10,000 SAMPLES FROM A STANDARD TRIANGULAR DISTRIBUTION.

N	n	PARAMETRIC	METHOD					
			GRAPHICAL					
			$P_R = P(E(X_R))$			$P_R = E(P(X_R))$		
			LS-X	LS-Z	GMFR	LS-X	LS-Z	GMFR
8	4	0.076	0.076	0.066	0.071	0.063	0.054	0.058
	8	0.072	0.073	0.058	0.065	0.057	0.044	0.050
15	4	0.063	0.063	0.057	0.060	0.053	0.047	0.050
	8	0.062	0.063	0.054	0.058	0.053	0.044	0.049
	15	0.061	0.062	0.052	0.057	0.052	0.042	0.048
30	4	0.055	0.056	0.054	0.055	0.048	0.046	0.047
	8	0.055	0.056	0.051	0.053	0.049	0.044	0.047
	15	0.055	0.056	0.051	0.053	0.050	0.044	0.047
	30	0.055	0.056	0.051	0.053	0.050	0.044	0.047

lowest at n=4 and did not vary among the n by more than 3%. There is thus no substantial advantage, with respect to precision, to using n>4.

Another factor in the selection of the subsample size is the possibility of skewness. Therefore, simulations were also conducted using a skewed triangular distribution for which the mode was 20%, rather than 50%, of the distance from the lower to the upper limit. For all sample sizes, using n=N resulted in $\overline{P}_5$<0.02 and, for N=15 and N=30, using n=N/2 resulted in $\overline{P}_5$ being about 0.03. Using n=4, however, resulted in $\overline{P}_5$ being between 0.04 and 0.05 for all N. Using n=4 also resulted in substantially lower standard deviations for $\hat{x}_5$ than using n=N/2 and n=N.

Because for a nonskewed population the smallest subsample size considered performed little, if any, worse than larger subsamples, and because the possibility and consequences of skewness give good reason to restrict the subsample size, a subsample size n=4 is recommended. Limited consideration was given to a smaller subsample (n=2), but for N=8 this resulted in substantial bias in $\overline{P}_5$ ($\overline{P}_5$=7.0%) and in a 20% increase in the standard deviation of $\hat{x}_5$; also, the use of so few data markedly increases the sensitivity of results to an occasional unusually low datum.

Consideration was also given to the possible effects of nonrandom sampling by determining the bias introduced into $\overline{P}_5$ if the method recommended above is used when samples are not obtained randomly. Two nonrandom sampling schemes were investigated. First, samples were taken in an entirely systematic fashion highly correlated with variate, data being uniformly distributed over percentiles with the i^{th} datum (X_i) being set equal to x_{p_i}, where p_i=(i-0.5)/N. With such a scheme, $P(\hat{x}_5)$ equals 0.015, 0.024, and 0.034 for N equal to 8, 15,

and 30, respectively. Although $P(\hat{x}_5)$ is therefore substantially biased, this is an extremely unrealistic depiction of the sampling and the biases are extreme upper limits. Sampling schemes with a more realistic systematic component would result in much lower biases.

The second nonrandom sampling scheme used was stratified sampling in which each member of the sample was assumed to be randomly sampled from a restricted percentile range. The range for the ith datum of a sample was $[(100\%)(i-0.5)/N] \pm 25\%$; i.e., a fifty percentile range centered on the value used for the i^{th} datum of the systematic sample discussed above. For low and high i, the range was compressed so that percentiles were maintained between 0 and 100 and so that, over the entire sample, each member of the population had an equal chance of being drawn. Specifically, for low i, if a percentile was computed by the above formula to be <0, its absolute value was used. This results in the ranges for low i being narrower than the nominal 50% (as small as 27% for i=1 and N=30) and sampling within the ranges being somewhat skewed to low percentiles; analogous compression and skewing occurred for high i. Because the recommended procedure would heavily employ data with low i, the systematic component of this procedure is therefore even greater than implied by the restriction of sampling to fifty percentile ranges. Using this sampling scheme, ten-thousand samples of size 8, 15, and 30 from a standard triangular distribution were computer generated and FAVs were calculated from each sample by the procedure recommended above based on random sampling. $\overline{P_5}$ was 0.040, 0.042, and 0.044 for N equal to 8, 15, and 30, respectively. This small bias suggests that, even with a strong systematic element in the actual sampling, an assumption of random sampling performs well as long as there is also a substantial random element in the sampling, or at least an element that is not correlated with variate.

APPLICATION OF RECOMMENDED PROCEDURE AND ALTERNATIVES TO DATA SETS

In the previous sections all issues necessary for the recommendation of a procedure for FAV calculation have been considered. The recommended new procedure assumes the set of ln(MAV)s is a random sample from a triangular distribution with unknown location and scale parameters. The mechanics of the procedure can be summarized as follows:

> The ln(MAV)s are ranked and each assigned a cumulative probability $P_R=R/(N+1)$, where R is the rank and N the number of data in the set. A line of the form $\ln(MAV)=\hat{S}\cdot\sqrt{P_R}+\hat{L}$ is fit to the four points with P_R nearest 0.05. (The square root of P_R constitutes transformation to the variate of a standard triangular distribution somewhat different than, but equally valid to, that used above; it is used here because it is simpler to calculate when $P_R<0.5$). $\hat{S}$, $\hat{L}$, and FAV are calculated as follows:

$$\hat{S} = \sqrt{\frac{\sum^4(\ln MAV)^2-(\sum^4 \ln MAV)^2/4}{\sum^4 P_R-(\sum^4\sqrt{P_R})^2/4}}$$

$$\hat{L} = (\sum^4 \ln MAV-\hat{S}\cdot\sum^4\sqrt{P_R})/4$$

$$\ln FAV = \hat{S}\cdot\sqrt{0.05}+\hat{L}$$

> Example calculations are provided in Appendix 1.

This procedure was applied to the example data sets in Tables 1 and 2. The FAVs thus calculated for each data set are included in Tables 7 and 8, along with the lowest MAV as a reference.

Also included in Tables 7 and 8 are FAVs calculated for each data set using the old procedure presented in the November 28, 1980 version of the Guidelines (1). This procedure can be described as follows:

> The ln(MAV)s are ranked and assigned to fixed intervals with width = 0.25 and with the first interval starting at the lowest ln(MAV). Each interval is assigned a cumulative proportion $P=R_{max}/N$, where N is the number of ln(MAV)s and R_{max} is the rank of the largest ln(MAV) in the interval. Each interval is assigned a variate V=average ln(MAV) within the interval. The nonempty interval with the highest P less than or equal to 0.05 (or the

TABLE 7. FAVs CALCULATED FROM SMAV DATA SETS BY OLD PROCEDURE, RECOMMENDED NEW PROCEDURE, AND VARIOUS MODIFICATIONS OF RECOMMENDED NEW PROCEDURE.

MATERIAL	WATER	N	LOWEST SMAV	OLD PROCEDURE	NEW PROCEDURE	MODIFICATIONS OF RECOMMENDED NEW PROCEDURE						
						n=N/2	n=N	UNIFORM DIST.	NORMAL DIST.	SLOPE CHANGE	PARAM. METHOD	NONRANDOM SAMPLING
COPPER	FRESH	45	0.23	0.29	0.34	0.39	0.22**	0.33	0.35	0.35	0.33	0.39
DDT	FRESH	42	0.36	1.12	0.87	0.85	0.49	0.81	0.91	0.87	0.79	1.28
CADMIUM	SALT	31	41	59	62	60	87	59	63	62	66	80
CADMIUM	FRESH	29	0.020	0.024	0.024	0.017	0.065	0.023	0.025	0.025	0.031	0.030
TOXAPHENE	FRESH	29	1.30	1.58	1.59	1.47	1.37	1.57	1.61	1.61	1.63	1.97
ZINC	FRESH	29	8.9	7.0	9.7	10.1	11.4	9.4	9.9	10.0	11.7	9.1
ENDRIN	FRESH	28	0.15	0.18	0.20	0.22	0.07**	0.20	0.21	0.21	0.21	0.30
MERCURY	SALT	26	3.5	3.7	3.9	2.6	2.6*	3.9	4.0	4.0	4.1	4.6
ZINC	SALT	24	166	173	195	166	108*	195	195	201	200	252
LINDANE	FRESH	22	2.0	2.2	2.6	4.3	5.8	2.7	2.6	2.8	3.3	5.8
COPPER	SALT	22	28	23	27	25	25	28	27	28	30	30
NICKEL	FRESH	22	54	56	55	73	99	56	55	57	69	69
DIELDRIN	SALT	21	0.70	0.71	0.67	0.72	0.92	0.68	0.66	0.67	0.80	0.82
ALDRIN	FRESH	21	4.0	3.0	3.9	3.2	0.8**	3.9	3.9	3.9	4.3	4.3
ENDRIN	SALT	21	0.037	0.037	0.034	0.030	0.027*	0.035	0.034	0.036	0.040	0.043
HEPTACHLOR	SALT	19	0.057	0.053	0.080	0.133	0.155	0.090	0.077	0.100	0.097	0.223
DIELDRIN	FRESH	19	2.5	2.5	2.7	2.5	1.3*	2.8	2.6	2.7	2.8	3.4
LINDANE	SALT	19	0.17	0.16	0.26	0.51	0.54	0.30	0.24	0.36	0.33	0.92
CHROMIUM(VI)	SALT	19	2000	1260	1770	1750	1200*	1820	1750	1830	2060	2010
CHROMIUM(III)	FRESH	18	33	32	29	38	41	31	28	30	37	39
HEPTACHLOR	FRESH	18	0.90	0.52	0.77	0.52	0.96	0.81	0.75	0.79	0.93	0.99
NICKEL	SALT	17	152	137	149	106	78*	161	144	151	170	204
DDT	SALT	17	0.140	0.121	0.147	0.167	0.108	0.158	0.142	0.155	0.161	0.212
ALDRIN	SALT	16	1.50	1.34	1.53	1.91	1.76	1.68	1.47	1.63	1.70	2.16
CYANIDE	FRESH	15	57	52	56	60	27**	58	54	57	59	64
TOXAPHENE	SALT	14	0.110	0.070	0.094	0.107	0.058	0.117	0.084	0.102	0.120	0.161
CHROMIUM(VI)	FRESH	14	67	21	56	115	493	88	46	79	89	175
CHLORDANE	FRESH	14	3.0	2.4	2.1	2.8	4.7	2.6	1.9	2.1	2.8	3.6
SELENIUM	FRESH	13	340	263	200	201	441	254	178	205	275	391
SELENIUM	SALT	13	600	410	480	430	310*	510	470	500	550	600
ENDOSULFAN	SALT	12	0.040	0.034	0.034	0.029	0.004**	0.039	0.032	0.035	0.038	0.044

Table 7. Continued

MATERIAL	WATER	N	LOWEST SMAV	OLD PROCEDURE	NEW PROCEDURE	MODIFICATIONS OF RECOMMENDED NEW PROCEDURE					
						n=N/2	n=N	UNIFORM DIST.	NORMAL DIST.	SLOPE CHANGE	PARAM. METHOD	NONRANDOM SAMPLING
ARSENIC(III)	FRESH	12	810	440	340	260	620	430	310	430	560	820
MERCURY	FRESH	11	5.0	3.7	2.6	1.6	2.9	3.5	2.3	2.7	3.6	5.2
SILVER	FRESH	10	0.0019	0.0014	0.0014	0.0017	0.0003*	0.0019	0.0012	0.0018	0.0018	0.0019
SILVER	SALT	10	4.7	3.3	3.3	3.9	2.8	4.4	3.0	3.7	4.0	4.7
ENDOSULFAN	FRESH	10	0.340	0.218	0.183	0.214	0.125*	0.258	0.158	0.186	0.251	0.340
CHLORDANE	SALT	8	0.400	0.090	0.200	0.200	0.378	0.352	0.162	0.278	0.313	0.280
GEOMETRIC MEAN OF RATIOS OF FAV BY MODIFIED PROCEDURE TO THAT BY RECOMMENDED PROCEDURE:				0.89	1.00	1.04	0.93	1.11	0.96	1.07	1.19	1.48
NUMBER OF DATA SETS FOR WHICH FAV BY MODIFIED PROCEDURE DIFFERS FROM THAT BY RECOMMENDED PROCEDURE BY MORE THAN A FACTOR OF 1.4:				5	-	8	26	2	0	0	3	13
NUMBER OF DATA SETS FOR WHICH FAV BY MODIFIED PROCEDURE DIFFERS FROM THAT BY RECOMMENDED PROCEDURE BY MORE THAN A FACTOR OF 2.0:				2	-	1	12	0	0	0	0	6

Table 8. FAVs CALCULATED FROM FMAV DATA SETS BY OLD PROCEDURE, RECOMMENDED NEW PROCEDURE, AND VARIOUS MODIFICATIONS OF RECOMMENDED NEW PROCEDURE.

MATERIAL	WATER	N	LOWEST FMAV	OLD PROCEDURE	NEW PROCEDURE	MODIFICATIONS OF RECOMMENDED NEW PROCEDURE						
						n=N/2	n=N	UNIFORM DIST.	NORMAL DIST.	SLOPE CHANGE	PARAM. METHOD	NONRANDOM SAMPLING
CADMIUM	SALT	25	75	45	70	79	119	69	70	73	97	77
COPPER	FRESH	23	0.30	0.34	0.38	0.45	0.26	0.39	0.38	0.39	0.42	0.58
MERCURY	SALT	23	3.5	3.7	3.8	2.5	2.7*	3.8	3.8	3.8	4.0	4.4
DDT	FRESH	20	1.30	1.22	1.28	0.98	0.55**	1.30	1.27	1.29	1.37	1.46
ZINC	SALT	20	166	166	182	146	109*	187	180	187	192	236
CADMIUM	FRESH	18	0.048	0.038	0.058	0.086	0.234	0.069	0.054	0.063	0.075	0.143
ENDRIN	FRESH	17	0.44	0.33	0.40	0.38	0.09**	0.41	0.40	0.42	0.44	0.46
COPPER	SALT	17	28	27	25	23	26	27	25	26	28	32
CHROMIUM(VI)	SALT	17	2490	1940	2370	2130	1550*	2440	2340	2370	2540	2680
DIELDRIN	SALT	16	0.70	0.67	0.53	0.55	0.76	0.58	0.51	0.55	0.68	0.77
ENDRIN	SALT	16	0.037	0.036	0.031	0.021	0.017*	0.032	0.030	0.032	0.036	0.041
HEPTACHLOR	SALT	16	0.057	0.044	0.061	0.106	0.117	0.076	0.055	0.073	0.077	0.148
LINDANE	SALT	16	0.170	0.121	0.192	0.398	0.395	0.248	0.170	0.272	0.263	0.578
NICKEL	FRESH	16	65	50	66	93	122	75	62	70	76	103
ZINC	FRESH	15	13.7	10.4	12.3	13.7	19.1	14.1	11.5	12.7	14.2	19.3
ALDRIN	FRESH	14	7.4	4.0	6.7	5.8	1.1**	6.9	6.6	6.8	7.2	7.5
DDT	SALT	14	0.140	0.102	0.130	0.149	0.092	0.149	0.122	0.138	0.148	0.182
NICKEL	SALT	14	310	160	210	130	110*	240	200	220	270	320
CHROMIUM(III)	FRESH	13	33	30	23	28	33	27	21	24	31	36
ALDRIN	SALT	13	3.7	3.5	3.3	3.2	2.3*	3.5	3.2	3.4	3.5	3.9
TOXAPHENE	SALT	13	0.110	0.065	0.087	0.094	0.047	0.112	0.077	0.095	0.113	0.147
DIELDRIN	FRESH	12	4.5	1.6	3.7	2.7	1.0**	3.9	3.6	3.8	4.1	4.6
TOXAPHENE	FRESH	12	1.30	0.99	1.07	1.12	0.87	1.24	1.00	1.08	1.20	1.42
SELENIUM	SALT	12	600	440	440	330	320*	490	420	470	540	600
ENDOSULFAN	SALT	11	0.040	0.033	0.033	0.028	0.003**	0.039	0.031	0.034	0.036	0.042
HEPTACHLOR	FRESH	10	1.00	0.75	0.50	0.34	0.52	0.68	0.44	0.53	0.67	1.00
CHROMIUM(VI)	FRESH	10	67	8	23	38	243	56	16	31	39	67
SELENIUM	FRESH	10	340	154	167	209	513	267	136	181	256	340
LINDANE	FRESH	10	10.0	8.2	6.4	7.0	7.0	8.1	5.8	6.9	7.6	10.0
CYANIDE	FRESH	10	77	58	63	60	25**	68	61	64	71	77
SILVER	SALT	10	4.7	3.3	3.3	3.9	2.8	4.4	3.0	3.7	4.0	4.7
SILVER	FRESH	9	0.0019	0.0011	0.0013	0.0013	0.0003*	0.0018	0.0011	0.0016	0.0017	0.0017

Table 8. Continued

MATERIAL	WATER	N	LOWEST FMAV	OLD PROCEDURE	NEW PROCEDURE	MODIFICATIONS TO RECOMMENDED NEW PROCEDURE						
						n=N/2	n=N	UNIFORM DIST.	NORMAL DIST.	SLOPE CHANGE	PARAM. METHOD	NONRANDOM SAMPLING
MERCURY	FRESH	9	5.00	3.42	0.94	0.94	2.32	1.79	0.73	1.12	1.89	4.76
ENDOSULFAN	FRESH	9	0.340	0.208	0.169	0.169	0.094*	0.248	0.144	0.172	0.234	0.319
ARSENIC(III)	FRESH	8	880	570	220	220	730	410	170	250	450	790
CHLORDANE	FRESH	8	6.3	3.7	4.0	4.0	4.6	5.3	3.6	4.0	4.6	5.6
CHLORDANE	SALT	8	0.400	0.090	0.200	0.200	0.378	0.352	0.162	0.278	0.313	0.280
GEOMETRIC MEAN OF RATIOS OF FAV BY MODIFIED PROCEDURE TO THAT BY RECOMMENDED PROCEDURE:				0.91	1.00	1.01	0.89	1.23	0.92	1.08	1.25	1.58
NUMBER OF DATA SETS FOR WHICH FAV BY MODIFIED PROCEDURE DIFFERS FROM THAT BY RECOMMENDED PROCEDURE BY MORE THAN A FACTOR OF 1.4:				9	-	8	29	7	1	1	5	17
NUMBER OF DATA SETS FOR WHICH FAV BY MODIFIED PROCEDURE DIFFERS FROM THAT BY RECOMMENDED PROCEDURE BY MORE THAN A FACTOR OF 2.0:				5	-	1	13	1	0	0	2	7

interval with the lowest P if no interval has a P less than 0.05) is designated Interval A. The next highest nonempty interval is designated Interval B. The FAV is then computed as $\ln(FAV)=V_A+(V_B-V_A)/(P_B-P_A)\cdot(0.05-P_A)$.

Criticisms of this procedure include:

(1) The formula P=R/N is positively biased as discussed earlier and thus results in negative bias in the FAV.

(2) The positive bias in the cumulative proportions is increased by using the maximum rank in an interval rather than the average rank, when there is more than one ln(MAV) in the interval.

(3) Often only one ln(MAV) is in Interval A or Interval B or both, making the method quite sensitive to data variation.

(4) A linear relationship of P versus V is assumed, which is equivalent to assuming a rectangular distribution; as discussed earlier this is contraindicated by the available data sets.

(5) The use of intervals is meant to cause pooling of ln(MAV)s which are indistinguishable; the interval width of 0.25 was selected because it is a typical value for the standard deviation of replicate acute toxicity tests; this value may not be appropriate for all species and materials and is strictly appropriate only when ln(MAV)s are based on only one toxicity test. More importantly, this pooling method works effectively only for Interval A, because the starting point for Interval B is fixed by that for Interval A and therefore does not necessarily properly pool data in the vicinity of Interval B. In any event, this pooling serves no useful purpose except to prevent the slope for interpolations and extrapolations from being inappropriately calculated based on two identical, or nearly identical, ln(MAV)s, a purpose which can be better served by routinely using more points to assess data trends.

(6) The use of intervals containing variable numbers of ln(MAV)s makes the method sensitive to minor changes in the data set which may move ln(MAV)s into or out of intervals; this sensitivity can be quite marked and can even be anomalous. For example, in the SMAV data set for heptachlor in fresh water (Table 1), Interval A would contain the lowest two SMAVs and Interval B would consist of the third lowest SMAV. However, if the lowest SMAV was just 6% lower (for example, because a new toxicity test for that species lowered the mean slightly), the two lowest SMAVs would be sufficiently separated to be in separate intervals, which would become the new, and markedly different, Intervals A and B. The calculated FAV would change from 0.52 to 0.83, a change that not only is much larger than the change in the SMAV that caused it (60% versus 6%), but also is in the opposite direction to the change in the SMAV.

These criticisms are sufficient to warrant the replacement of this old procedure with the new procedure recommended above. However, the two procedures generally do not produce markedly different results (Tables 7 and 8). On the average, FAVs calculated using the old procedure are only 11% lower than those calculated using the new procedure for the SMAV data sets and 9% lower for the FMAV data sets. Individual FAVs were within a factor of 1.4 for over 75% of the FMAV data sets and over 85% of the SMAV data sets and within a factor of 2.0 for over 85% of the FMAV sets and 94% of SMAV data sets. Where differences are greater than twofold, comparison of the FAVs with the data sets does not clearly indicate that one or the other of these procedures results in more questionable FAVs.

The consequences of modifying the major features of the recommended new procedure were also explored to determine how sensitive FAVs are to such changes. If the sensitivity is low, any objection to compromises or approximations used in arriving at the recommended new procedure are largely irrelevant, because more exact analysis or different compromises (within reason) would have little effect in practice. If sensitivity is high, the basis for the recommended new procedure becomes more critical and further examination is warranted.

Three features of the recommended new procedure were modified to span the range over which they could reasonably be varied. The assumed distribution was changed to rectangular and to normal. The subset size (n) was changed to N/2 and N. The percentile estimation method was changed to the graphical method with slope formula LS-X, but still with $P_R=E(P(X_R))$, and to the parametric method (best linear unbiased estimate). The FAVs for these modifications are included in Tables 7 and 8. Also included in these tables are (a) the geometric mean of the ratios of the FAV by each modified procedure to that by the recommended procedure, (b) the number of data sets for which the FAV by each

modified procedure differs by more than a factor of 1.4 from that by the recommended procedure, and (c) the number of data sets for which the FAV by each modified procedure differs by more than a factor of 2.0 from that by the recommended procedure.

Changes in the assumed distribution, in the percentile estimation method, and in the subset size to N/2 had only minor effects on results. The geometric means of the ratios of the FAVs by these modifications to that by the recommended procedure were close to 1.0 (0.92-1.25). For individual data sets, FAVs by these modifications differed by more than a factor of 1.4 from the FAVs by the recommended new procedure for no more than 20% of the data sets and by more than a factor of 2.0 for no more than 6% of the data sets.

Modification of subset size to n=N, however, caused major changes. The geometric means of the ratios of the FAV by this modification to that by the recommended procedure were close to 1.0 (0.93 for SMAVs, 0.89 for FMAVs), but individual FAVs changed by more than a factor of 1.4 for over 70% of the SMAV data sets and for nearly 80% of the FMAV data sets and by more than a factor of 2.0 for about one-third of both the SMAV and FMAV data sets. Such differences do not demonstrate, per se, that this modified procedure is less appropriate than the recommended new procedure, but it does raise such a suspicion. In particular, when n=N, an unusually large number of sets have a FAV that is well below both the lowest MAV and the FAV calculated by the recommended new procedure. Using the location and scale parameter estimates from the modified procedure with n=N, the fiducial probability that the lowest MAV could be so high was evaluated for each data set; where this probability is <0.20 a single asterisk is placed next to the FAV for n=N and where the probability is <0.10 a double asterisk is used. The frequency of these marked entries suggests that

using n=N is in fact inappropriate. This is directly related to the existence of statistically significant skewness in the data sets, positive skewness resulting in inappropriately low FAVs when n=N and negative skewness resulting in high FAVs.

A related observation that also contraindicates the use of n=N is that, when compared to the recommended new procedure and the diverse modifications already mentioned, the modification with n=N both frequently produces the lowest FAV (for 20 SMAV data sets and 19 FMAV data sets) and frequently produces the highest FAV (for 14 SMAV data sets and 12 FMAV data sets). Such frequent occupation of both extremes is again due to the variable skewness of the sets and is indicative of the error of assuming all data are equally useful in estimating low percentiles when distributional assumptions are not completely met. Furthermore, these problems with using n=N are not restricted to the modification with n=N already discussed. Graphical methods with the other slope formulas and other distributional assumptions (e.g., normal) were tested using n=N with similar results. Likewise, the best linear unbiased method using n=N and assuming a normal distribution showed similar problems. (This latter method is equivalent to the simple approach of calculating a sample mean and unbiased standard deviation (12) and estimating the fifth percentile as lying 1.645 standard deviations below the mean, -1.645 being the fifth percentile of a standard normal distribution.)

Another advantage of not using n=N is that certain semiquantitative data can be used. Acute tests on some materials with some species produce 'greater than' values because concentrations high enough to cause effects were not, or could not be used (because of solubility or time constraints). Regardless of the reason, because such data are usually for resistant species, they can

usually be used if n=4, but, if n=N, either they must be excluded, thereby biasing the data set, or additional acute tests must be conducted, thereby increasing costs. Finally, it should be noted that the use of n=4 does not constitute 'not using all the data', because all data are used in setting ranks and cumulative probabilities and thus in selecting which four data will be used explicitly in final calculations; rather, the use of n=4 is more properly interpreted as a simple scheme of giving greater weight to those MAVs which provide the most information about the fifth percentile.

The consequences of nonrandom sampling can also be partly addressed here. Assuming that the available data sets somehow resulted from the strictly systematic sampling scheme discussed earlier, FAVs were calculated by assigning cumulative probabilities $P_R=(R-0.5)/N$ to the ranked ln(MAV)s and interpolating between the two data with P_R nearest 0.05 (or extrapolating using the lowest two points if N<10), the interpolation being based on the assumption of a triangular distribution. The results of this exercise are included in the last columns of Tables 7 and 8 and indicate that higher FAVs are produced than by the recommended new procedure, but the differences average only about a factor of 1.5 for SMAV sets and 1.6 for FMAV sets and are less than a factor of 1.4 for 65% of the SMAV sets and 55% of the FMAV sets. Considering that this alternative sampling scheme is so extreme, and that therefore a scheme with a more realistic systematic component would produce results much nearer those obtained by the recommended new procedure, this further suggests that the issue of the sampling assumption is not of great importance.

A final point that should be emphasized is that, whether SMAVs or FMAVs are used, the same conclusions are reached regarding the appropriate attributes of

the procedure for estimating fifth percentiles. Also, although it does not bear on the recommendations made here, it is interesting to note that FAVs calculated from SMAVs are similar to those calculated from FMAVs. The geometric mean of the ratios of the FAV computed from FMAVs to that computed from SMAVs, by the recommended procedure, was 1.04. The two FAVs differ by a factor of 1.4 for only 12 sets, by a factor of 2.0 for only 6 sets, and by more than a factor of 2.8 for no sets.

DISCUSSION

The recommended new procedure uses linear extrapolation or interpolation to estimate the fifth percentile of a statistical population of mean acute values (MAVs) from which the available MAVs are assumed to have been randomly obtained. The available MAVs are ranked from low to high and the cumulative probability for each is calculated as $P_R=R/(N+1)$, where R = rank and N = number of MAVs in the set. Extrapolation or interpolation is based on an assumed linear relationship between $\sqrt{P_R}$ and ln(MAV), and uses only the four points with P_R closest to 0.05 because this subset provides the most useful information concerning the fifth percentile.

The bases for the new procedure are mostly mathematical, with some input from toxicological and practical considerations. The FAV, however, is basically a toxicological value, and the acceptability of any calculation procedure to toxicologists will be based on the acceptability of the resulting FAVs. Most aquatic toxicologists will judge the acceptability of an FAV by comparing it with the lowest MAVs in the data set, and thus it is quite appropriate that the four MAVs with estimated cumulative probabilities closest to 0.05 be given the most weight in calculating the FAV; in fact, the recommended procedure is largely a formalization of the way one would obtain a FAV by 'eyeballing' the data.

An important property of the new procedure is that the resulting FAV is not very sensitive to modifications in the procedure or slight changes in the data set. A variety of calculation procedures all produced FAVs that were quite similar for most data sets. In addition, the recommended new procedure rarely produced either the highest or lowest of the FAVs obtained with the procedures examined. Another important property of a procedure is its performance with

data sets which contain apparent discontinuities in the lower tail. For heptachlor, lindane, and chlordane in salt water and chromium(VI) in fresh water, the lowest SMAV is at least a factor of 10 lower than the second lowest SMAV (Table 1). In addition, for these four and cadmium in fresh water, the lowest FMAV is at least a factor of 10 lower than the second lowest FMAV (Table 2). Of these, only chromium(VI) in fresh water has a very large range of FAVs in Tables 7 and 8. Even though the two lowest MAVs are far apart, most of these data sets seem to provide adequate information about the FAV because similar FAVs were obtained using a variety of procedures. These examples support the idea that the best approach to take toward calculating the FAV is to select a procedure that is best on the average and then use it with all data sets, except possibly in extraordinary cases.

An unfortunate aspect of the methodology for calculating the FAV is the necessity of extrapolating to estimate the 0.05 cumulative probability for small data sets; if extrapolations become too great, the FAVs will be suspect. For only 5 of the 74 data sets is the FAV more than a factor of 2 lower than the lowest MAV. Thus, for the available data sets this procedure rarely extrapolates much below the lowest value in the data set.

Overall, the recommended new procedure is the best of the procedures examined, regardless of whether the FAV is calculated from SMAVs or FMAVs. It is a straightforward procedure for interpolation or extrapolation based on fitting a line to the most useful points. It produces results similar to and usually intermediate to those obtained by other reasonable procedures. In addition, the calculations are relatively easy to perform with the aid of a hand calculator, as described in Appendix 1. The major weakness of this procedure is

that it assumes the same degree of tailing for all data sets. Fortunately, for most data sets the FAV is not very dependent on the assumed degree of tailing and deviation from the assumed intermediate degree of tailing is not too critical. Other procedures would suffer as much or more from the same or other weaknesses.

REFERENCES

1. U.S. EPA. 1980. Federal Register. 45:79318-79379. November 28.

2. Javitz, H. and J. Skurnick. 1980. Analyses of relationships among various aquatic toxicity test results. Final Report, Task 9, EPA Contract 68-01-3887.

3. Heit, M. 1981. Letter to J. H. McCormick. U.S. EPA, Duluth, MN.

4. Sokal, R.R. and F.J. Rohlf. 1969. Biometry. Freeman, San Francisco.

5. Lloyd, E.H. 1952. Least squares estimation of location and scale parameters using order statistics. Biometrika, 39:88-95.

6. Johnson, N.L. and S. Kotz. 1970. Distributions in statistics: Continuous unvariate distributions - 1. Wiley, New York.

7. Box, G.E.P., W.G. Hunter, and J.S. Hunter. 1979. Statistics for Experimenters. Wiley, New York.

8. Cunnane, C. 1978. Unbiased plotting positions - a review. J. Hydrol. 37:205-232.

9. Sprent, P. and G.R. Dolby. 1980. Response to query. Biometrics, 36:547-550.

10. Barker, F., Y.C. Soh, and R.J. Evans. 1988. Properties of the geometric mean functional relationship. Biometrics, 44:279-281.

11. David, F.N. and N.L. Johnson. 1954. Statistical treatment of censored data. Part I: Fundamental formulae. Biometrika, 41:228-240.

12. Zar, J.H. 1974. Biostatistical Analyses. Prentice-Hall, Englewood Cliffs, New Jersey.

APPENDIX 1

A. General Instructions for Recommended New Procedure for FAV Calculation.

1. Based on data set size (N), determine four ranks (R) with cumulative probabilities (P_R=R/(N+1)) closest to 0.05; for N<60, this will be R=1 through 4; for 60<N<80, R=2 through 5; for 80<N<100, R=3 through 6; etc.

2. From the data set select the four MAVs with the desired ranks and calculate P_R for each of these MAVs.

3. Fit a line to ln(MAV) vs $\sqrt{P_R}$ using the following equations for slope ($\hat{S}$) and intercept ($\hat{L}$) and calculate the FAV:

$$\hat{S} = \sqrt{\frac{\sum^4 (\ln MAV)^2 - (\sum^4 \ln MAV)^2/4}{\sum^4 P_R - (\sum^4 \sqrt{P_R})^2/4}}$$

$$\hat{L} = (\sum^4 \ln MAV - \hat{S} \cdot \sum^4 \sqrt{P_R})/4$$

$$A = \ln FAV = \hat{S} \cdot \sqrt{0.05} + \hat{L}$$

$$FAV = e^A$$

B. Example Calculation for Chlordane in Salt Water (N=8).

Rank	MAV	lnMAV	$(\ln MAV)^2$	P_R=R/(N+1)	$\sqrt{P_R}$
4	6.4	1.8563	3.4458	0.44444	0.66667
3	6.2	1.8245	3.3290	0.33333	0.57735
2	4.8	1.5686	2.4606	0.22222	0.47140
1	0.4	-0.9163	0.8396	0.11111	0.33333
Sum:		4.3331	10.0750	1.11110	2.04875

$$\hat{S} = \sqrt{\frac{10.0750 - (4.3331)^2/4}{1.11110 - (2.04875)^2/4}} = 9.3346$$

$$\hat{L} = [4.3331 - (9.3346)(2.04875)]/4 = -3.6978$$

$$A = (9.3346)\ (\sqrt{0.05}) - 3.6978 = -1.6105$$

$$FAV = e^{-1.6105} = 0.1998$$

C. Example Computer Program in BASIC Language for Calculating the FAV

```
10  REM THIS PROGRAM CALCULATES THE FAV WHEN THERE ARE LESS THAN
20  REM 59 MAVS IN THE DATA SET.
30  X=0
40  X2=0
50  Y=0
60  Y2=0
70  PRINT "HOW MANY MAVS ARE IN THE DATA SET?"
80  INPUT N
90  PRINT "WHAT ARE THE FOUR LOWEST MAVS?"
100 FOR R=1 TO 4
110 INPUT V
120 X=X+LOG(V)
130 X2=X2+(LOG(V))*(LOG(V))
140 P=R/(N+1)
150 Y2=Y2+P
160 Y=Y+SQR(P)
170 NEXT R
180 S=SQR((X2-X*X/4)/(Y2-Y*Y/4))
190 L=(X-S*Y)/4
200 A=S*SQR(0.05)+L
210 F=EXP(A)
220 PRINT "FAV = "F
230 END
```

D. Example Printout from Program

```
HOW MANY MAVS ARE IN THE DATA SET?
? 8
WHAT ARE THE FOUR LOWEST MAVS?
? 6.4
? 6.2
? 4.8
? .4
FAV = 0.1998
```