Weed Science. 1986. Volume 34:485–489

Accuracy and Precision of Analyses for 2,4-D and Picloram in Streamwater by Ten Contract Laboratories[1]

LOGAN A. NORRIS[2]

Abstract. **In a test of laboratory precision and accuracy, streamwater samples fortified with 0, 10, or 50 μg ae/L 2,4-D [(2,4-dichlorophenoxy)acetic acid] and picloram (4-amino-3,5,6-trichloro-2-pyridinecarboxylic acid) were submitted in quadruplicate for residue analysis to 10 laboratories that analyze samples for a fee. Varying amounts of the herbicides were detected by eight laboratories, and two reported no residues. Analyses from most laboratories had a negative bias (ranging from −2 to −92%), although a positive bias appeared in a few instances. Precision was highly variable, the coefficient of variation ranging from 3.4 to 200%. The accuracy range in 19 of 20 cases was ±100 to ±78% for 10 and 50 μg/L 2,4-D and ±100 to ±114% for 10 and 50 μg/L picloram, respectively. Results show that careful selection of contract laboratories and a quality assurance program should be part of monitoring for water contamination with herbicides.**

Additional index words. **Water contamination, herbicide residues, quality assurance.**

INTRODUCTION

Monitoring for herbicides in forest and rangeland waters has become an important part of brush control programs, particularly on public lands. The herbicides 2,4-D and picloram, alone or in combination, accounted for about 75% of all herbicide use in 1982 in national forests (12). Water samples are often analyzed on a contract basis by public and private laboratories, but as Corneliussen (5) stated, results "may vary considerably from the true values".

This study determined the accuracy and precision with which analyses for the two common herbicides in water were made by 10 contract laboratories selected at random. Although Edwards et al. (7) conducted a similar study of the quality of general nutrient analysis, this study differs from many interlaboratory tests with pesticides (1, 4), in that a common procedure was not required and participants did not know they were being tested. The purpose of the test was to determine the accuracy and precision of analyses likely to be obtained by the typical forest manager, who may lack the training or experience necessary for selecting laboratories or evaluating results.

METHODS

Care was taken that the water samples were not identified as test samples, so that they would be processed like others analyzed for a fee. I indicated through another person that they were from a program being conducted by a large public agency. All analyses were paid for. Correspondence and purchase orders identified only the USDA Forest Service and gave only a street address. No research group or individual working with chemical residues was identified.

Requests for a price quotation and a description of the analytical method used for measuring 2,4-D and picloram in water were sent to 76 laboratories (mostly in the western United States) believed to analyze herbicides for a fee. Twenty of the 44 laboratories replying were able to analyze water for 2,4-D and picloram by gas chromatography with a sensitivity of at least 1 μg/L. From these, 10 laboratories were selected at random. Each was sent a purchase order (based on the price quotation) and a set of 16 water samples for analysis.

Preparation of water samples. Clean 3-L glass containers were filled with 2 L of streamwater from an undisturbed forest area. To each container was added 2 or 10 ml water containing a specified amount of herbicide as the sodium salt. After 8 g NaOH was added as a preservative, the container was tightly capped and inverted repeatedly to ensure mixing. The samples were of four types, each with different concentrations of herbicides (Table 1).

Because upstream operations may contaminate authentic samples and interfere with analysis of known chemicals, and because dicamba (3,6-dichloro-2-methoxybenzoic acid) is occasionally used in forestry and may be carried along with 2,4-D and picloram in the analyses, it was added as an "unknown" to achieve a concentration of 10 or 50 μg ae/L in sample types III and IV, respectively. The laboratories were told the samples were collected in connection with the application of 2,4-D and picloram, with no mention of dicamba.

A total of 224 samples were prepared and assembled into sets of 16, each set containing four samples of each type numbered individually so that it appeared there were 16 different samples rather than four samples in quadruplicate. Each of the 10 contract laboratories and each of three public agency laboratories (control laboratories) analyzed one set of samples, and one set was saved for reference and possible replacement. The control laboratories knew some samples were fortified with 2,4-D, picloram, and an unknown contaminant, but they did not know the concentrations of herbicide or which samples were fortified. Their purpose was to determine, independently, if the herbicides in the samples were measurable.

Analysis of data. The mean, percent bias [(M-T)100 ÷ T, where M is the measured value and T is the true value],

[1] Received for publication May 27, 1985, and in revised form October 15, 1985.

[2] At the time that the research was conducted, Chief Res. Chem., For. Sci. Lab., Pac. Northwest For. Range Exp. Stn., U.S. Dep. Agric., For. Serv., Corvallis, OR 97331; now Head, Dep. For. Sci., Coll. For., Oregon State Univ., Corvallis, OR 97331.

Table 1. Herbicide concentrations in streamwater samples submitted for analysis.

Sample type	Herbicide concentration		
	2,4-D	Picloram	Dicamba
	(μg ae/L)		
I	0	0	0
II	10	50	0
III	50	10	10
IV	0	0	50

standard deviation, coefficient of variation and total error [$(E + 2\sigma)100 \div T$, where E is the mean error, σ is the standard deviation, and T is the true value] were calculated. The chi-square test (8) was applied to the analytical results for each chemical in each sample set from each laboratory. In the standard chi-square test of a hypothesized variance,

$$X^2(n)df = \frac{\sum_{i=1}^{N} (x_i - \mu_i)^2}{\sigma^2}$$

where X^2 = the chi-square statistic, n = the number of observations, df = degrees of freedom, σ^2 = hypothesized variance (i.e., the required level of accuracy, where $\sigma^2 = \frac{E^2}{\tau}$, E is the deviation allowed from the true value, and τ is the standard normal deviate for the specified probability level), x_i = value of the i^{th} observational unit estimated by a given laboratory, and μ_i = true value of the i^{th} unit. This test is normally used to determine if results of repli-

Table 2. Results of analysis by control and contract laboratories of water sample type II (10 μg/L 2,4-D and 50 μg/L picloram.).

Source of analysis	Herbicide	Sample number				Mean	Standard deviation	Bias[a]	Coefficient of variation	Total error[b]
		3	9	12	16	(μg/L)		(%)		
Control laboratories										
A	2,4-D	7.4	7.4	7.3	7.6	7.4	±0.1	−26	1.3	28
	Picloram	27	29	27	30	28.3	±1.5	−43	5.3	49
B	2,4-D	9.4	9.9	10	9.7	9.8	±0.3	−2	3.1	8
	Picloram	51	51	49	49	50	±1.2	0	2.4	5
C	2,4-D	8.6	7.4	9.3	8.6	8.5	±0.8	−15	9.4	31
	Picloram	36	18	31	40	31.3	±9.6	−37	31	76
Averages				2,4-D		8.6	0.4	−14	4.6	22
				Picloram		36	4.1	−27	13	43
Contract laboratories[c]										
2	2,4-D	7.0	6.8	12	10	9.0	2.5	−10	28	60
	Picloram	13	16	44	37	27	15	−46	56	100
3	2,4-D	0[d]	0	0	0	0				
	Picloram	24	41	24	21	27	9.1	−46	33	81
4	2,4-D	11	14	11	11	12	1.3	+20	11	43
	Picloram	3.9	3.8	6.0	1.9	3.9	1.7	−92	44	99
6	2,4-D	21	24	22	21	22	1.4	+120	6.4	310
	Picloram	85	170	100	110	120	37	+140	32	280
7	2,4-D	0	8.0	8.0	7.0	5.8	3.9	−42	67	120
	Picloram	0	39	3.0	16	14	18	−72	113	140
8	2,4-D	7.8	9.7	9.3	8.4	8.8	0.9	−12	10	30
	Picloram	120	1500	370	380	592	620	+1084	100	3600
9	2,4-D	10	7.0	7.0	8.0	8.0	1.4	−20	17	48
	Picloram	41	30	25	38	33	7.3	−34	22	62
10	2,4-D	11	0	7.7	7.4	7.0	4.7	−30	72	130
	Picloram	64	0	45	21	32	28	−36	86	150
Averages				2,4-D		7.2	6.7	−28	93	160
				Picloram		85	181	+70	210	800

[a] $(M-T)100 \div T$, where M is the measured value and T is the true value.

[b] $(E + 2\sigma)100 \div T$, where E is the mean error, σ is the standard deviation, and T is the true value.

[c] No herbicide was detected in any sample by contract laboratories 1 and 5.

[d] 0 means reported as not detected.

Table 3. Results of analysis by control and contract laboratories of water sample type III (50 μg/L 2,4-D, 10 μg/L picloram, and 10 μg/L dicamba.).

Source of analysis	Herbicide	Sample number 2	6	7	10	Mean	Standard deviation	Bias[a]	Coefficient of variation	Total error[b]
		(μg/L)						(%)		
Control laboratories										
A	2,4-D	31	30	31	30	30	±0.6	−39	2	41
	Picloram	7.4	6.4	7.2	6.4	6.9	±0.5	−31	7.2	41
B	2,4-D	50	51	47	49	49	±1.7	−1	3.4	8
	Picloram	9.5	10	9.7	10	9.8	±0.2	−2	2	6
C	2,4-D	39	44	43	47	43	±3.3	−13	7.6	27
	Picloram	5.3	3.3	5.9	7.4	5.5	±1.7	−45	31	79
Averages				2,4-D		41	1.9	−18	4.3	25
				Picloram		7.4	0.8	−26	13	42
Contract laboratories[c]										
2	2,4-D	45	45	42	48	45	2.5	−10	5.5	20
	Picloram	2.6	4.2	3.7	16	6.6	6.3	−34	95	160
3	2,4-D	52	49	52	49	50	1.7	0	3.4	8
	Picloram	0[d]	0	0	8	2	4	−80	200	160
4	2,4-D	53	52	53	44	50	4.4	0	8.7	19
	Picloram	1.4	2.3	2.7	0	1.6	1.2	−84	75	110
6	2,4-D	0	130	90	80	75	54	+50	72	270
	Picloram	0	15	7	8	7.5	6.1	−35	81	150
7	2,4-D	0	37	33	26	24	17	−52	69	120
	Picloram	0	4	1	3	2	1.8	−80	90	120
8	2,4-D	61	44	38	48	48	9.7	−4	20	43
	Picloram	160	270	380	780	400	270	+3900	68	9300
9	2,4-D	22	48	44	40	38	11	−24	30	67
	Picloram	7	7	14	11	9.8	3.4	−2	34	70
10	2,4-D	19	21	12	35	22	9.6	−56	44	95
	Picloram	4	4.9	T	9.8	4.7	4	−53	85	130
Averages				2,4-D		35	24	−29	67	140
				Picloram		43	120	+330	290	2800

[a] $(M-T)100 \div T$, where M is the measured value and T is the true value.

[b] $(E + 2\sigma)100 \div T$, where E is the mean error, σ is the standard deviation, and T is the true value.

[c] No herbicide was detected in any sample by contract laboratories 1 and 5.

[d] 0 means reported as not detected.

cated analyses meet a predetermined level of "accuracy" at the 95% confidence level. (This definition of accuracy includes both accuracy and precision.) (8). In this study, I solved the formula for σ^2 and calculated the range of values that would encompass the "accuracy" likely to be attained in 19 of 20 trials. These results are expressed as the deviation from the true value which is likely to include the results of 19 of 20 replicated analyses of streamwater fortified with herbicides as in this study. Incorrectness may also be due to a consistent bias resulting from low recovery, inaccurate standards, or some other systematic error. I used Freese's modified chi-square test to evaluate the range of accuracy with bias removed (8).

RESULTS

The interval between the time that laboratories received samples and reported results varied from 4 to 49 days (average 20 days), and the per-sample charge for analysis varied from $50 to $110 (average $80).

Sample types I and IV. One-half of the contract laboratories reported false positives in at least one sample of types I and IV (data not presented); however, the magnitude of error was small. The highest value was 4.6 μg/L, and only 50% of the values were greater than 1 μg/L. None of the contract laboratories reported false positives in all samples. None of the contract laboratories confused dicamba with 2,4-D or picloram nor did they report the presence of any compound other than 2,4-D or picloram. One control laboratory reported a possible false positive, but the level was below the limit of quantification. All the control laboratories detected and reported dicamba, but they were aware an unknown contaminant would be present in some samples (data not presented).

Sample types II and III. Control laboratories found 2,4-D and picloram in all type II and III samples. Within and among

contract laboratories, the accuracy and precision of analysis of sample types II and III varied widely (Tables 2 and 3). The contract laboratories reported 29% false negatives, about equally distributed between 2,4-D and picloram. Laboratories 1 and 5 failed to detect either herbicide in any sample, accounting for two-thirds of all false negatives reported. Laboratory 3 reported most of the rest of the false negatives, failing to detect 10 μg/L 2,4-D in any sample and 10 μg/L picloram in three of four samples. None of the control laboratories reported false negatives.

Most laboratories reported less than the amount of herbicide added (negative bias). Positive bias was infrequent. Laboratory 8 consistently overestimated the concentration of picloram by a substantial margin. Precision was highly variable, the coefficients of variation ranging from 3.4 to 73% for 2,4-D and from 22 to 200% for picloram. Total error was smaller in analyses for 2,4-D than for picloram, but the averages are dominated by the high positive bias of laboratory 8. Control laboratory B was uniquely successful in both accuracy and precision.

The results are also expressed as the deviation from the true value which will encompass 19 of 20 replicated analyses (Table 4). An example may help illustrate what this means. Of 20 samples of water containing 10 μg/L 2,4-D and 50 μg/L picloram sent to laboratory 2, 19 of 20 times the results would fall between 6.9 and 13.1 μg/L 2,4-D and between 17 and 83 μg/L picloram (no correction for bias). Correction of the data for bias, using Freese's modified chi-square, markedly improved the performance of two control laboratories and of some contract laboratories. The great improvement in analysis for picloram when data were so corrected clearly illustrates the value of including fortified controls with samples in order to determine the percentage of recovery.

As a group, the control laboratories performed better than the contract laboratories. This may be attributable to inherently better analytical capability, or, where the differences in performance are smaller, to the fact that control laboratories knew of the test and were therefore more careful in applying standard procedures.

Table 4. Deviation from the true value (with and without correction for bias) which will encompass 19 of 20 replicated analyses of stream-water fortified with herbicides.

Source of analysis	2,4-D[a]				Picloram[a]			
	10 μg/L		50 μg/L		10 μg/L		50 μg/L	
	a	b	a	b	a	b	a	b
	(μg/L)							
Control laboratories								
A	3.3	0.3	25	0.7	4.1	0.6	27.7	1.8
B	0.4	0.3	2.1	2.1	0.4	0.3	1.3	1.4
C	2.1	1	9.3	4	6.1	2.1	26.1	1.6
Contract laboratories[b]								
2	3.1	3	6.9	3	8.2	7.6	33	19
3	. . .[c]	. . .	2	2.1	11	4.9	30	11
4	2.6	1.6	4.9	5.3	11	1.5	59	2
6	15	1.7	68	66	7.5	7.5	94	45
7	6.9	4.7	38	20	10	2.2	49	22
8	1.8	1	11	12	580	330	970	750
9	3	1.7	19	14	3.8	4.1	22	8.9
10	6.8	5.7	37	12	8.1	4.9	38	34

[a]Columns "a" show uncorrected values based on Tables 1 and 2, P = 0.95, 4 degrees of freedom; columns "b" show values corrected for bias, P = 0.95, 3 degrees of freedom (8).

[b]No herbicide was detected by laboratories 1 and 5; the exact accuracy level is undefined.

[c]No 2,4-D was detected at 10 μg/L by laboratory 3; the exact accuracy level is undefined.

DISCUSSION

The accuracy and precision of analysis in many laboratories was surprisingly poor, a result of some combination of poor laboratory procedure, inadequate training, or lack of experience. All can be corrected if the laboratory manager or analyst is aware of the problem. Horwitz (10), Gunther (9), and McFarren et al. (11) provide excellent discussions of the reliability of analytical results and the interpretation of pesticide residue data. Gunther (9) provides guidelines for the range within which measured pesticide residue values may be expected to vary, even in competent laboratories. At concentrations greater than 1 ppm (w/v), variation of ±10% is expected; at 0.1 ppm (w/v), ±20%; at 0.01 ppm (w/v), ±50%; and at 0.001 ppm (w/v), ±200% (usually more) (9). The results from most of the laboratories in this study did not meet these guidelines.

All laboratories should monitor their precision and accuracy routinely. Analysis of a replicated series of fortified samples, which will provide data on both accuracy and precision and allow appropriate corrections to be made, should be part of standard quality assurance procedures. When problems such as those shown by laboratories 1, 5, and 8 develop, they can then be corrected before erroneous results are reported. Quality control procedures are well developed and readily available for pesticide analyses (3, 13).

Resource managers and others with regulatory responsibilities would have been substantially misled by results from several of the laboratories included in this study. The study also shows the danger of relying on results of a single analysis of a critical sample. Laboratory selection and quality control are important parts of any herbicide monitoring program, and qualification procedures may be needed to assist in laboratory selection (6). Burke (2) cites four papers that include detailed descriptions of laboratory performance studies. Variations of the procedures used in these types of studies could be used for laboratory certification or prequalification before awarding a contract for analyses. The procedures outlined are beyond the capabilities of most land managers, but staff or consulting chemists could assist in the selection process, particularly when large numbers of samples are to be analyzed.

Edwards et al. (6) point out that, while science and art are important elements in pesticide analyses, craftmanship is essential. They urge inclusion of a quality control clause

in contracts for analytical services, but they note that such a clause is useless without a technique for evaluation and enforcement. Fortified samples of the type used in this test are one possible tool for this purpose.

ACKNOWLEDGMENTS

The assistance of Carroll Burkett, John Hazard, and Donald Henshaw, Pac. Northwest For. and Range Exp. Stn., U.S. Dep. Agric. For. Serv., in preparing and shipping samples, and providing statistical planning and analysis is gratefully acknowledged. This is Paper 1961, For. Res. Lab., Oregon State Univ. The research was supported through the National Agric. Pest. Impact Assessment Program, U.S. Dep. Agric. Serv., and Cooperative Aid Agreement PNW-191, with Oregon State Univ.

LITERATURE CITED

1. Aspila, K. I., J. M. Carron, and A.S.Y. Chan. 1977. Interlaboratory quality control study of the analysis of water for pesticides. J. Assoc. Off. Anal. Chem. 60:1097–1104.
2. Burke, J. A. 1978. The interlaboratory study in pesticide residue analysis. Pages 633–642 *in* H. Geissbuhler, ed. Advances in Pesticide Science, Vol. III, Part 3. Pergamon Press, Elmsford, NY.
3. Burke, J. A. and P. E. Corneliussen. 1975. Quality assurance in the Food and Drug Administration's pesticide residue analytical laboratories. Pages 25–31 *in* P. Koivistoinen and J. Koivurinta, eds. Environmental Quality and Safety Supplement. Vol. 3. Thieme, Stuttgart.
4. Burse, V. W., L. L. Needham, M. P. Korver, C. R. Lapeza, Jr., J. A. Liddle, and D. D. Bayse. 1983. Assessment of methods to determine PCB levels in blood serum: Interlaboratory study. J. Assoc. Off. Anal. Chem. 66:40–45.
5. Corneliussen, P. E. 1970. Pesticide residues in total diet samples (v), Pestic. Monit. J. 4:89–97.
6. Edwards, R. R., R. A. Dailey, and H. Cruse. 1975. Quality assurance in water analysis contracts. J. Am. Water Works Assoc. 67:363–366.
7. Edwards, R. R., D. L. Schilling, Jr., and T. L. Rossmiller. 1977. A performance evaluation of certified water analysis laboratories. J. Water Pollut. Control Fed. 49:1704–1712.
8. Freese, F. 1960. Testing accuracy. For. Sci. 6:139–145.
9. Gunther, F. A. 1980. Interpreting pesticide residue data at the analytical level. Residue Rev. 76:155–171.
10. Horwitz, W. 1981. Analytical measurements: How do you know your results are right? Adv. Chem. 160:411–438.
11. McFarren, E. F., R. J. Lishka, and J. H. Parker. 1970. Criterion for judging acceptability of analytical methods. Anal. Chem. 42:358–365.
12. Norris, L. A., H. W. Lorz, and S. V. Gregory. 1983. Forest chemicals, U.S. Dep. Agric., For. Serv., Pac. Northwest For. Range Exp. Stn., Gen. Tech. Rep. PNW-149, Portland, OR. 95 pp.
13. Sherma, J. 1981. Manual of analytical quality control for pesticides and related compounds in human and environmental samples. U.S. Environ. Prot. Agency, EPA 600/2-81-059. Research Triangle Park, NC.