

NWRI CONTRIBUTION 88-59

**UPPER GREAT LAKES CONNECTING CHANNELS
INTERLABORATORY PERFORMANCE EVALUATION STUDY
QM-12: TOTAL PHENOL IN WATER
FINAL REPORT**

by

W.C. Li, H.B. Lee and E. Kokotich

**Research and Applications Branch
National Water Research Institute
Canada Centre for Inland Waters
867 Lakeshore Road, P.O. Box 5050
Burlington, Ontario, Canada L7R 4A6**

and the Quality Management Work Group

December 1987

**UPPER GREAT LAKES CONNECTING CHANNELS
INTERLABORATORY PERFORMANCE EVALUATION STUDY**

QM-12: TOTAL PHENOL IN WATER

FINAL REPORT

by

W.C. Li, H.B. Lee and E. Kokotich

**Research and Applications Branch
National Water Research Institute
Canada Centre for Inland Waters
867 Lakeshore Road, P.O. Box 5050
Burlington, Ontario, Canada L7R 4A6**

December 1987

and

**The Quality Management Work Group
*sent to the QMWG for review and approval***

MANAGEMENT PERSPECTIVE

The Upper Great Lakes Connecting Channels (UGLCC) have been designated as "Areas of Concern" by the International Joint Commission. A Canada - U.S. binational study, involving the identification and assessment of the environmental impacts of toxic substances, in those areas, was initiated in 1984. In order to assist analytical laboratories, which are contributing data to the UGLCC study, to generate reliable and accurate data during the study, a Quality Management Work Group was formed and 13 interlaboratory performance evaluation studies were implemented.

This report summarizes and evaluates the results from the twelfth interlaboratory performance evaluation study, QM-12, which consisted of the analysis of total phenol in water. Results were received from four Canadian and three U.S. laboratories out of 10 participants. Overall, 81% of the data, received from the participants, were satisfactory and comparable. All participating laboratories have been provided with appropriate feedback.

Dr. J. Lawrence
Director
Research and Applications Branch

PERSPECTIVES DE GESTION

Les "Upper Great Lakes Connecting Channels" (UGLCC) ont été désignés comme "Zones à risques" par la Commission mixte internationale. Une étude binationale Canada - États-Unis, a été mise en oeuvre en 1984 pour caractériser et évaluer les effets environnementaux de substances toxiques dans ces zones. Pour aider les laboratoires d'analyses, qui contribuent à l'étude UGLCC, à obtenir des données sûres et précises dans le cadre de l'étude, un Groupe de travail en gestion qualitative a été formé, et 13 études d'évaluation de la performance inter-laboratoires ont été entreprises.

Le présent rapport résume et évalue les résultats de la douzième étude sur l'évaluation de la performance inter-laboratoires, QM-12, qui concernait l'analyse du phénol total dans l'eau. Sur 10 participants, des résultats ont été obtenus de quatre laboratoires canadiens et de trois laboratoires des États-Unis. Dans l'ensemble, 81 % des résultats reçus des participants étaient satisfaisants et comparables. Tous les laboratoires participants ont bénéficié d'une rétroaction très efficace.

D^r J. Lawrence
Directeur
Direction de la recherche et des
applications

ABSTRACT

The Upper Great Lakes Connecting Channels (UGLCC) recognizes Quality Assurance/Quality Control (QA/QC) aspects as crucial elements to the overall utility of study results. As part of the QA/QC program, thirteen interlaboratory performance evaluation studies were designed and conducted by the Quality Management Work Group.

This report describes the results from the twelfth interlaboratory performance evaluation study, QM-12, which consisted of the analysis of total phenol in water. Results were received from seven out of ten participating laboratories (four Canadian, three U.S.).

Data were evaluated for bias by Youden's ranking technique and results which deviated significantly from the median were flagged. The interlaboratory comparability of total phenol in water was satisfactory. There was good agreement between the interlaboratory medians and the design values. Included in this report is a summary of each laboratory's performance.

RÉSUMÉ

Les "Upper Great Lakes Connecting Channels" (UGLCC) reconnaissent les aspects Assurance de la qualité/Contrôle de la qualité (AQ/CQ) comme des éléments cruciaux de l'utilité globale des résultats de l'étude. Dans le cadre du programme AQ/CQ, treize études d'évaluation de la performance inter-laboratoires ont été conçues et effectuées par le Groupe de travail en gestion de la qualité.

Le présent rapport décrit les résultats de la douzième étude d'évaluation de la performance inter-laboratoires QM-12, qui portait sur l'analyse du phénol total dans l'eau. On a obtenu les résultats de sept sur dix laboratoires participants (quatre canadiens, trois des États-Unis).

Les données ont été évaluées pour les biais par la technique de classement de Youden, et les résultats qui s'écartaient sensiblement de la médiane ont été marqués. La comparabilité inter-laboratoires était satisfaisante pour le phénol total dans l'eau. Il y avait une bonne corrélation entre les médianes inter-laboratoires et les valeurs théoriques. Le rapport présente un résumé du rendement obtenu par chaque laboratoire.

1.0 INTRODUCTION

The Upper Great Lakes Connecting Channels (UGLCC) have been designated as "Areas of Concern" by the International Joint Commission (IJC). To identify and deal with the environmental problems, a three year, binational study was started in 1984, involving Canadian and U.S. environmental and resource agencies, to study the St. Marys, St. Clair and Detroit Rivers, and Lake St. Clair. The study involves identifying, quantifying and determining the environmental impacts of conventional and toxic substances from various sources.

The UGLCCS recognizes Quality Assurance/Quality Control (QA/QC) aspects as crucial elements to the overall utility of study results. As part of the QA/QC program, thirteen interlaboratory performance evaluation (QC) studies were designed and conducted by the Quality Management Work Group. The goal of these QC studies is to assist analytical laboratories, which are producing data for the UGLCC study, to generate reliable, accurate data and to assess their overall performance during this study. A total of some 100 parameters (organic, inorganic and physical properties) in three types of matrices (water, sediment and biota), will be assessed.

This twelfth interlaboratory study, QM-12, was initiated on April 2, 1986. It involved the analysis of total phenol in water. The original deadline for reporting results was set for May 30, 1986. However, since several laboratories were late in reporting, the study was not closed until October 10, 1986.

2.0 STUDY PROFILE

From the returned questionnaires, the following 10 laboratories affirmed that they would participate in this study: U014, U049, U057, U077, U079, U089, U094, U063, U072 and U090. By the time the study was closed the last three laboratories had not sent back any results. See the list of participants at the end of this report.

Each laboratory was provided with four ampules and a one-litre sample of naturally occurring surface water. Two of the ampules were used to spike two 500 mL aliquots of the one-litre sample of naturally occurring surface water provided while the other two ampules were used to spike two 500 mL aliquots of the laboratory's own organic-free water. All ampules were well-characterized reference materials, prepared by the Quality Assurance Project Team, Research and Applications Branch of the National Water Research Institute (NWRI) and were stored at 4°C before distribution. The design values and interlaboratory medians for total phenol in waters are given in Table 2. The design values were verified by in-house and external analyses.

Participants were asked to analyze samples 1201 - 1204 for total phenol, using their in-house procedures and standards. In order to estimate the precision of such analysis, these samples were sent out in blind duplicate pairs, as shown in Table 1.

3.0 RESULTS AND DISCUSSION

3.1 Analytical Methodology

In this study, total phenol in water was analyzed by all participants with the colorimetric determination based on either the manual or automated 4-aminoantipyrine (4-AAP) method with distillation. For the sample pretreatment, five out of seven laboratories reporting results used manual distillation and two used automated distillation. For the method of detection, U014 used both automated and manual 4-AAP methods. While three laboratories used a manual 4-AAP method, the other three laboratories used an automated 4-AAP method. See Table 3 for the details of analytical methodologies.

3.2 Data Evaluation

All raw data submitted by the participants are listed in the data summary (Appendix II). Individual lab results for total phenol were evaluated by Youden's ranking technique (Ref. 1) for the detection of bias as well as a computerized flagging procedure (Ref. 2). A laboratory's results are judged biased high or low when its total rank is outside of a statistically allowable range. For a further explanation of the ranking and flagging procedures, see Appendix I. This statistical procedure, which semi-quantitatively evaluates data accuracy, is widely used in other interlaboratory QC studies. See Table 4 for a summary of total phenol data: ranking and flagging.

The accuracy of total phenol results is also summarized in Table 5. In this table, the number of results reported, the number of results flagged VH, H, L and VL were summed, and the percentages of results flagged were calculated. The statements of biased results are also included.

Paired sample plots are included as a graphical illustration of systematic vs random error as well as precision and accuracy of the participants' data (see Appendix III). The explanation of the detailed paired sample plots was given in the previous report (Ref. 3).

3.3 General Comments

Only three of the seven laboratories reported their data by the originally set deadline (U079, U089, U094). Lab U014 submitted two sets of total phenol results by using two different techniques. Accordingly, they were assigned lab numbers U014A and U014B. Computer printouts with raw data were sent to all reporting laboratories for verification on December 4, 1987. All laboratories returned their results verified. A final data summary was sent to the participating labs, the Quality Management Work Group, the Work Group Chairman and the MC and AIC Chairmen on January 23, 1987.

The overall interlaboratory performance of total phenol analysis in waters was satisfactory, except some of the data from U014B, U077 and U079. Interlaboratory medians agreed closely with the design values and the calculated recoveries ranged from 97.2 to 101%. After rejection of outliers, the interlaboratory RSD was better than 14%. The difference between the interlaboratory means and medians were less than 10% for all four water samples. The precision of within-lab duplicate analysis was rather good for all participants with the RSD better than 15%. The reported detection limit among the participants ranged from 1 to 5 µg/L.

According to the Youden plots, most results were precise and accurate with the exception of the following cases. Laboratory U014B's results were high for samples 1201 and 1204 and not precise for samples 1202 and 1203; laboratory U077 results were high and not precise for samples 1202 and 1203 and laboratory U079 results were high and not precise for samples 1201 and 1204. It indicated both random and systematic errors were present for the above-mentioned laboratories.

Although the results from laboratory U014 showed that the automated 4-AAP method (U014A) was more accurate and precise than the manual 4-AAP method (U014B), overall, both methods used by the other laboratories were generally comparable and accurate.

3.4 Lab-Specific Comments

Laboratory U014A's results were accurate and precise with no flags or bias statements. Precision was excellent with the RSD better than $\pm 2\%$. Laboratory U014B's results had two VH flags and one L flag. Sixty-three percent of their results were flagged. Precision was less satisfactory for samples 1202 and 1203 with the RSD $> \pm 12\%$.

Laboratory U049's results were accurate and precise with no flags or bias statements. Precision was better than $\pm 3\%$ RSD.

Laboratory U057's results were accurate and precise with no flags or bias statements. Precision was better than $\pm 7\%$ RSD.

Laboratory U077's results were satisfactory except for one VH flag on sample 1203. Twenty-five percent of results were flagged. Precision was less satisfactory with $\pm 14\%$ RSD.

Laboratory U079's results had one VH flag and one H flag. These results were biased high. Thirty-eight percent of results were flagged. Precision was less satisfactory for samples 1201 and 1204 with $\pm 20\%$ RSD.

Laboratory U089's results were accurate and precise with no flags or bias statements. Precision was excellent with the RSD better than $\pm 3\%$.

Laboratory U094's results were accurate and precise with no flags or bias statements. Precision was excellent with the RSD better than $\pm 1\%$.

ACKNOWLEDGEMENTS

The authors sincerely thank all participants for their cooperation, Wendy Horn, Richard Szawiola and Dallas Takeuchi for their assistance.

REFERENCES

1. Youden, W. J. and Steiner, E.H. Statistical Manual of AOAC, published by AOAC, P.O. Box 540, Benjamin Franklin Station, Washington, D.C. 20044 (1975).
2. Clark, J.L. Evaluation of Performance of Laboratories Determining Water Quality Constituents through Natural Water Samples whose True Values are Unknown. In summary of Conference Presentations. Envirometrics 81, pp. 54-55, 1981. Alexandria, Virginia, April 8-10, 1981.
3. Szawiola, R., Horn, W., Takeuchi, D. and the Quality Management Work Group. Upper Great Lakes Connecting Channels Interlaboratory Performance Evaluation Study QM-9: Total Mercury in Surface Water. NWRI Contribution No. 87-129, October 1987.

LIST OF PARTICIPANTS

Barringer Magenta Ltd., Rexdale, Ontario
Beak Analytical Services, Mississauga, Ontario
Detroit Wastewater Treatment Plant, Analytical Laboratory, Detroit,
Michigan
Michigan Department of Natural Resources, Lansing, Michigan
Ontario Ministry of the Environment, London, Ontario
Ontario Ministry of the Environment, Water Quality, Rexdale, Ontario
U.S. Geological Survey - NWQL, Arvada, Colorado

The following laboratories were given samples, but did not submit any results:

Michigan Department of Public Health, Lansing, Michigan
U.S. Army Corps of Engineers - Environmental Analysis Branch, Detroit,
Michigan
Zenon Environmental Inc., Burlington, Ontario

TABLE 1

Samples Distributed For Analysis In QM-12

Samples	Description
1201	In-house sample #4, used to spike 500 mL aliquot of a naturally occurring surface water provided
1202	In-house sample #5, used to spike organic-free laboratory water
1203	Same as 1202
1204	Same as 1201

TABLE 2

Design values and interlaboratory medians for total phenol in waters

All values are in $\mu\text{g/L}$.

Parameter	Sample Number			Sample Number		
	<u>1201 and 1204</u>			<u>1202 and 1203</u>		
	Design	<u>Median</u>		Design	<u>Median</u>	
	Value	1201	1204	Value	1202	1203
Total Phenol	12.0	12.1	12.0	57.6	56.5	56.0

TABLE 3

Analytical Methodology for Total Phenol in Water

Lab No.	Sample Pretreatment	Method of Detection
U014A	Manual Distillation	Automated 4-aminoantipyrine (4-AAP) method
U014B	Manual Distillation	Manual 4-AAP method
U049	Manual Distillation	Manual 4-AAP method
U057	Manual Distillation	Manual 4-AAP method
U077	Manual Distillation	Manual 4-AAP method
U079	Manual Distillation	Automated 4-AAP method
U089	Automated Distillation	Automated 4-AAP method
U094	Automated Distillation	Automated 4-AAP method

OM12 TOTAL PHENOL IN WATER

LOWER LIMIT FOR USE OF BASIC ACCEPTABLE ERROR=12.00 BASIC ACCEPTABLE ERROR= 2.40 CONCENTRATION ERROR INCREMENT= .10
LABORATORIES YET TO REPORT: U063, U072, U090
LABORATORY RESULTS OMITTED ARE NONE

SAMPLE LAB NO	1201 REPORTED VALUE	1201 RANK	1202 REPORTED VALUE	1202 RANK	1203 REPORTED VALUE	1203 RANK	1204 REPORTED VALUE	1204 RANK
U049	10.	1.00	53.	1.00	55.	1.00	10.	1.50
U057	11.	2.00	50.	2.00	50.	2.00	10.	1.50
U077	14.	6.00	57.	5.00	70.	6.00	12.	4.50
U079	15.0	7.00	61.	6.00	62.0	7.00	20.5	6.00
U089	12.0	3.50	55.0	3.00	54.0	3.00	11.5	3.00
U094	12.2	3.50	58.	4.00	57.	3.00	12.2	3.00
U14A	12.	3.50	56.	4.00	58.	4.00	12.	3.00
U14B	10.5	6.00	58.	6.50	49.0	1.00	10.5	4.50
MEDIAN	12.100		56.500		56.000		12.000	
CONC.								

LAB NO.	TOTAL RANK	AVERAGE RANK	NO. OF SAMPLES RANKED	SUMMARY OF FLAGGING	DETECTION LIMIT
U049	8.50	2.125	4		2.115
U057	6.50	1.625	4		1.5
U077	30.00	7.500	4	VH	1.5
U079	12.50	3.125	4	MVH	1.5
U089	22.50	5.625	4		2.5
U094	18.00	4.500	4		
U14A	22.50	5.625	4	VHLVH	
U14B					
OVERALL AVERAGE RANK IS		4.500			

LAB NO.	TOTAL RANK	AVERAGE RANK	NO. OF SAMPLES RANKED	SUMMARY OF FLAGGING	DETECTION LIMIT
U057	6.50	1.625	4		1.5
U049	8.50	2.125	4		2.115
U089	12.50	3.125	4		1.5
U14A	18.00	4.500	4		2.5
U094	22.50	5.625	4		1.5
U14B	22.50	5.625	4		2.5
U077	30.00	7.500	4	VHLVH	
U079				VH	
OVERALL AVERAGE RANK IS		4.500			

TOTAL PHENOL

TABLE 5

Summary of Total Phenol Results by Lab Based on the Youden Ranking
Technique and Computerized Flaging Procedure

Lab No.	No of Results Reported	No. of Results Flagged				% Flagged	Comments
		VH	H	L	VL		
U014A	4	0	0	0	0	0	Satisfactory
U014B	4	2	0	1	0	63	Flagged VH on sample 1201, 1204 and L on 1203, these results are erratic.
U049	4	0	0	0	0	0	Satisfactory
U057	4	0	0	0	0	0	Satisfactory
U077	4	1	0	0	0	25	Satisfactory except for flagged VH on sample 1203.
U079	4	1	1	0	0	38	Flagged H on 1201 and VH on 1204. Ranking indicates results are biased high.
U089	4	0	0	0	0	0	Satisfactory
U094	4	0	0	0	0	0	Satisfactory

* Each H or L flag was counted as half a flag.

APPENDIX I

Glossary of Terms

APPENDIX I

Glossary of Terms(1) Ranking

Ranking is a non-parametric statistical technique used for the detection of pronounced systematic error (bias) in interlaboratory studies. According to Youden's procedure, rank 1 is given to the laboratory that provided the lowest result, rank 2 to the next lowest. In case of a tie, the average rank is given to the tied laboratories. Results with a < sign are not ranked. For each parameter, the total rank of each laboratory is the sum of individual ranks on each sample. In the case of six test samples and ten laboratories, the 5% probability limits for ranking scores are 14 and 52. A laboratory with a score lower than 14 is identified as biased low. Similarly, a laboratory with a total rank higher than 52 is biased high. In both cases, their results are classified as outliers. In cases where a laboratory did not provide all the results, or some of the results were not ranked, the average rank instead of total rank was used for the determination of biased statements.

The more comparable, i.e., better, laboratories should have ranks in the middle rather than at the extreme ends. However, laboratories with middle ranks do not necessarily mean that they provide more consistent results since very high results (high ranks) and very low results (low ranks) would average out to yield a total rank close to the median. Therefore, ranking alone is not sufficient to determine the performance of a laboratory.

(2) Flagging

When the true values of constituents in test samples are unknown, individual results can be evaluated in terms of their absolute differences from the interlaboratory medians. Medians are chosen rather than means since they are not influenced by a moderate number of extreme

values. By this flagging technique, all results are graded into the following three groups in the order of decreasing accuracy: (1) results with no flags, (2) results with H or L flags, and (3) results with VH or VL flags. Before evaluation is performed, three parameters, namely, Lower Limit for use of Basic Acceptable Error (LLBAE), Basic Acceptable Error (BAE), and Concentration Error Increment (CEI) are to be set. LLBAE is usually set at the lower end of the medians in the test samples. An 20% error at LLBAE is considered reasonable for total phenol and thus this is used as BAE. For samples whose medians are at or below LLBAE, the results are evaluated according to the following formulae:

	Absolute difference between sample and median results	\leq BAE	: acceptable
BAE <	Absolute difference between sample and median results	$\leq 1.5 \times$ BAE:	H or L
	Absolute difference between sample and median results	$> 1.5 \times$ BAE:	VH or VL

For samples whose medians are above the LLBAE, the allowable BAE is augmented by adding an increment to the BAE. This increment is calculated by multiplying the CEI by the difference between the sample median and LLBAE values. In this study, the CEI is set at 0.10. Sample results are again evaluated by the above three formulae except that the augmented BAE is used instead of BAE.

For further discussion on this evaluation technique, please refer to the original paper by Clark.

Bias: A set of results is said to be biased when the set exhibits a tendency to be either higher or lower than some standard - the standard which has been used in the analysis of our studies thus

far has been the performance of all other participating laboratories. The ranking procedure employed in testing for bias is described in W.J. Youden's paper, "Ranking Laboratories by Round-Robin Tests" from Precision Measurement and Calibration, H.H. Ku, Editor, NBS Special Publication 300 - Volume 1, U.S. Government Printing Office, Washington, D.C., 1969. In this paper, Youden establishes the rationale for evaluating laboratories' performance by ranking results. In our use of the procedure there is about one chance in twenty of deeming a set of results biased when in fact it is not, that is, $t = 0.05$.

Codes

W: A "W" code is used with a reported result when no measurement was possible due to no response of the instrument to the sample. The "W" is preceded by the smallest determinative division that can be used in the units used in reporting.

T: The "T" code is used with values between the Criterion of Detection and the "W" value. The Criterion of Detection is commonly thought of by many as the limit of detection.

H : high
 VH: very high
 L : low
 VL: very low

APPENDIX II

UGLCC Interlaboratory Performance Evaluation Study

QM-12: Total Phenol in Water

Final Data Summary

QM12 TOTAL PHENOL IN WATER

PRINTOUT PREPARED: 87/01/19.

PARAMETER: TOTAL PHENOL

UG/L

SAMPLE RESULTS

1201 1202 1203 1204

LAB

U049	10.	53.	55.	10.
U057	11.	50.	50.	10.
U077	14.	57.	70.	12.
U079	15.	61.	62.	20.
U089	12.0	55.0	54.0	11.5
U094	12.2	58.	57.	12.2
U14A	12.	56.	58.	12.
U14B	18.	58.	49.	18.

TOTAL LABS REPORTING	8	8	8	8
----------------------	---	---	---	---

TOTAL LABS USED	8	8	8	8
-----------------	---	---	---	---

MEAN	13.02500	56.00000	56.87500	13.21250
------	----------	----------	----------	----------

STD DEV	2.55329	3.38062	6.77047	3.71346
---------	---------	---------	---------	---------

MEDIAN	12.10000	56.50000	56.00000	12.00000
--------	----------	----------	----------	----------

DESIGN VALUE	12	57.6	57.6	12
--------------	----	------	------	----

APPENDIX III

YOUDEN'S TWO SAMPLE PLOTS

Legend for Youden Plots



Laboratories



D

Design value



M

Interlaboratory median

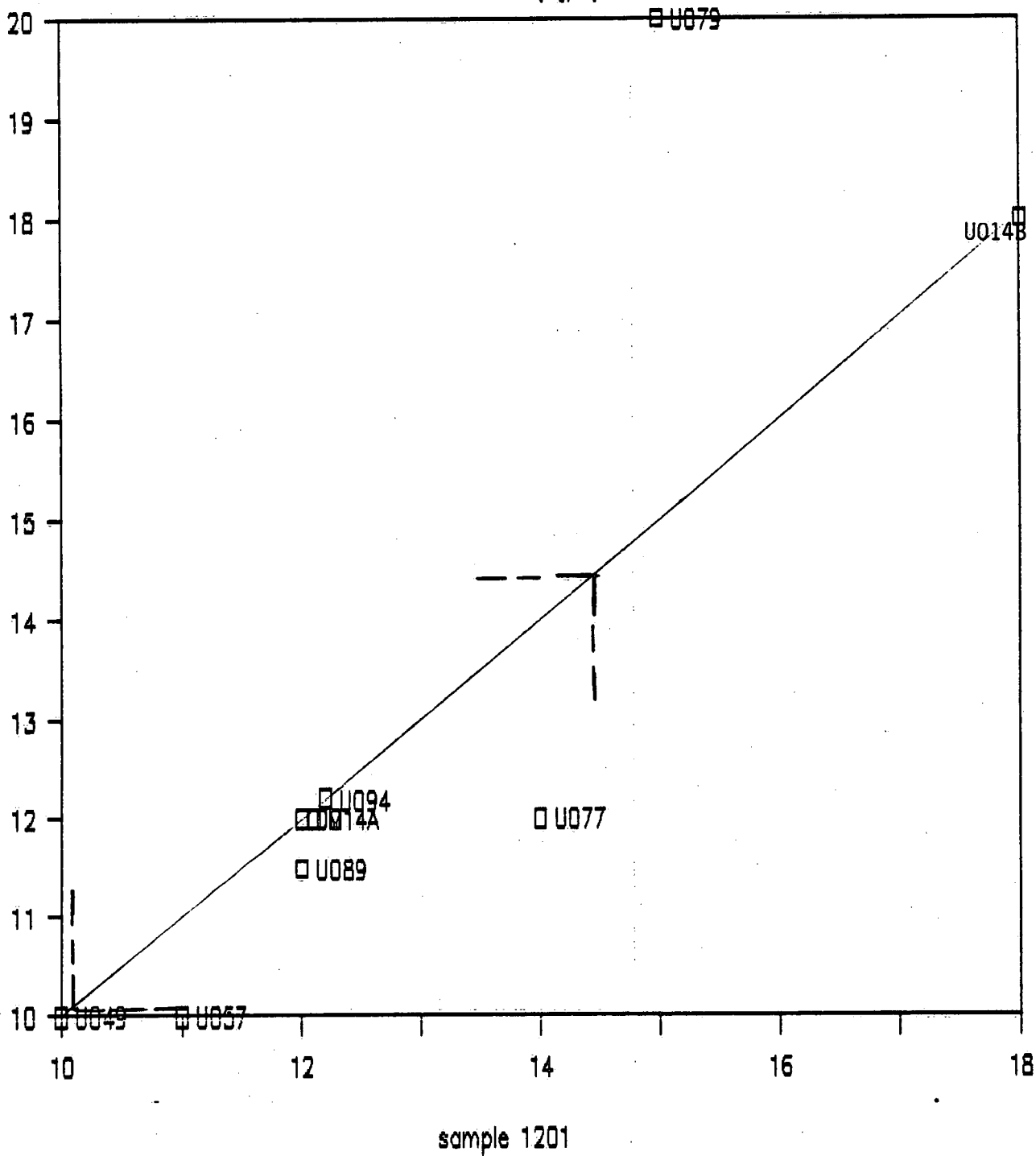


Range of the design value

PAIRED SAMPLE PLOT- 1201 & 1204

Total Phenol (ug/L)

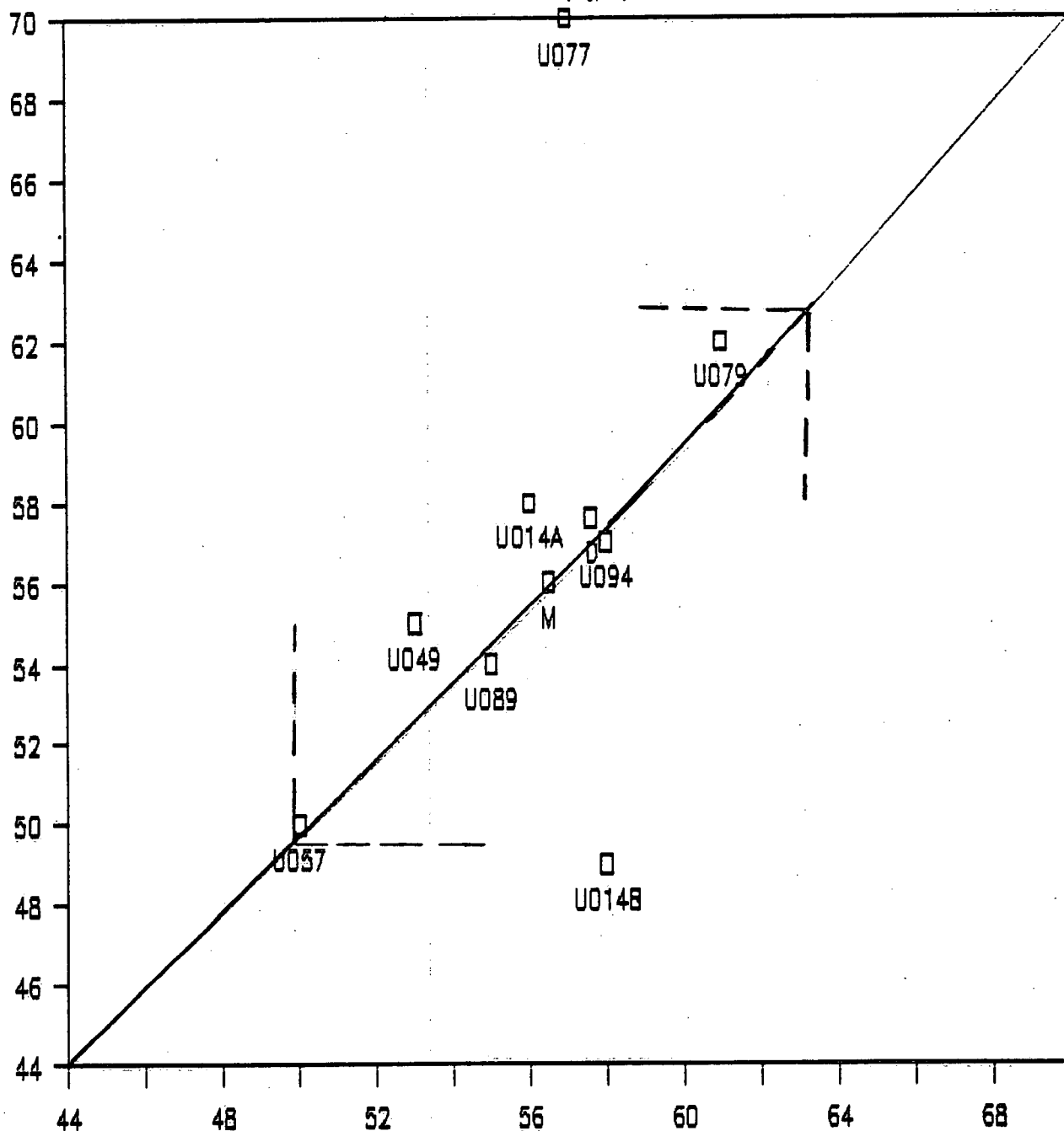
sample 1204



PAIRED SAMPLE PLOT- 1202 & 1203

Total Phenol (ug/L)

sample 1203



sample 1202