Station 4 JUL 09 2002

ENGINEERING DATA TRANSMITTAL

1. EDT 633969

S

2. To: (Receiving Organization) Distribution	3. From: (Originating Organization) Data Development & Interpretation	4. Related EDT No.: N/A
5. Proj./Prog./Dept./Div.: RPP/Engr./DD&I/Process Engr.	6. Design Authority/Design Agent/Cog. Engr.: M. R. Adams	7. Purchase Order No.: N/A
8. Originator Remarks: This document is being released into the supporting document system for retrievability purposes.		9. Equip./Component No.: N/A
		10. System/Bldg./Facility: N/A
		12. Major Assm. Dwg. No.: N/A
11. Receiver Remarks: N/A	11A. Design Baseline Document? ☐ Yes ☒ No	13. Permit/Permit Application No.: N/A
		14. Required Response Date: N/A

15. DATA TRANSMITTED

(A) Item No.	(B) Document/Drawing No.	(C) Sheet No.	(D) Rev. No.	(E) Title or Description of Data Transmitted	(F) Approval Designator	(G) Reason for Transmittal	(H) Originator Disposition	(I) Receiver Disposition
1	RPP-11352 1 JA 7/9/02	N/A	0	Statistical Optimization Study: Estimating Analyte Concentrations in Unsampled Waste Tanks	N/A	2	1	1

16. KEY

Approval Designator (F)	Reason for Transmittal (G)		Disposition (H) & (I)	
E, S, Q, D OR N/A (See WHC-CM-3-5, Sec. 12.7)	1. Approval 2. Release 3. Information	4. Review 5. Post-Review 6. Dist. (Receipt Acknow. Required)	1. Approved 2. Approved w/comment 3. Disapproved w/comment	4. Reviewed no/comment 5. Reviewed w/comment 6. Receipt acknowledged

17. SIGNATURE/DISTRIBUTION (See Approval Designator for required signatures)

(G) Reason	(H) Disp.	(J) Name	(K) Signature	(L) Date	(M) MSIN	(G) Reason	(H) Disp.	(J) Name	(K) Signature	(L) Date	(M) MSIN
		Design Authority						Steven R. Wilmarth		07/08/02	
		Design Agent									
2	1	Cog. Eng. M.R. Adams	MRAdams	07/03/02							
2	1	Cog. Mgr. J.G. Field	J G Field	07/08/02							
		QA									
		Safety									
		Env.									

18.	19.	20.	21. DOE APPROVAL (if required)
Mary Van Dorn M.E. Van Dorn 7/8/02 Signature of EDT Originator / Date	N/A Authorized Representative for Receiving Organization / Date	J G Field J.G. Field 7/08/02 Design Authority/ Cognizant Manager / Date	Ctrl No. ____ ☐ Approved ☐ Approved w/comments ☐ Disapproved w/comments

BD-7400-172-2 (10/97)

BD-7400-172-1

DISTRIBUTION SHEET

To	From	Page 1 of 1
Distribution	Data Development & Interpretation	Date 07/08/02
Project Title/Work Order		EDT No. EDT-633969
RPP-11352, Rev. 0, "Statistical Optimization Study: Estimating Analyte Concentrations in Unsampled Waste Tanks"		ECN No. N/A

Name	MSIN	Text With All Attach.	Text Only	Attach./ Appendix Only	EDT/ECN Only
CH2M HILL Hanford Group, Inc.					
M. R. Adams	R2-12	X			
J. G. Field	R2-12	X			
L. A. Fort	R2-11	X			
N. W. Kirch	R3-73	X			
D. C. Lowe	R2-58	X			
L. M. Sasaki	R2-12	X			
L. R. Webb	R1-10	H			
S. R. Wilmarth	R2-12	X			
Office of River Protection					
DOE Reading Room	H2-53	H			

A-6000-135 (10/97)

Statistical Optimization Study: Estimating Analyte Concentrations in Unsampled Waste Tanks

M. R. Adams and S. R. Wilmarth
CH2M HILL Hanford Group, Inc.
Richland, WA 99352
U.S. Department of Energy Contract DE-AC27-99RL14047

EDT/ECN: EDT-633969 UC: 2070
Cost Center: 7G300 Charge Code: 12214-74
B&R Code: EW 3120074 Total Pages: 18

Key Words:
Tank waste analytes, analyte concentrations, statistical estimation, optimization study

Abstract:
The proof-of-principle is demonstrated for the estimation of key analytes in unsampled waste tanks using statistical methods. A subset of analytes from liquid and solid double-shell tank wastes was used. Statistical methods compared the "between tank" and "cross tank" variability to determine the feasibility of utilizing statistics to estimate concentrations in unsampled tanks.

TRADEMARK DISCLAIMER. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof or its contractors or subcontractors.

Printed in the United States of America. To obtain copies of this document, contact: Document Control Services, P.O. Box 950, Mailstop H6-08, Richland WA 99352, Phone (509) 372-2420; Fax (509) 376-4989.

JUL 09 2002
DATE:
HANFORD RELEASE
STA: 4

7/8/02
Release Approval Date Release Stamp

Approved For Public Release

RPP-11351
Revision 0

Statistical Optimization Study
Estimating Analyte Concentrations in Unsampled Waste Tanks

M.R. Adams
S.R. Wilmarth
CH2M HILL Hanford Group, Inc.

Date Published
June 2002

CH2MHILL
Hanford Group, Inc.

Prepared for Office of River Protection

Contract #DE-AC06-99RL14047

Approved for public release; distribution is unlimited.

TABLE OF CONTENTS

1.0 INTRODUCTION AND BACKGROUND 1

2.0 SUMMARY OF STATISTICAL TECHNIQUES 2

3.0 STUDY RESULTS 3
3.1 DOUBLE-SHELL TANK LIQUID RESULTS 3
3.2 DOUBLE-SHELL TANK SOLIDS RESULTS 6

4.0 POTENTIAL BENEFICIAL APPLICATIONS OF THE RESULTS 10

5.0 RECOMMENDATIONS 11

APPENDICES

APPENDIX A: STATISTICAL METHODS EMPLOYED A-1

LIST OF TABLES

Table 3-1. Double-Shell Tank Liquid Variability Ratios 3

Table 3-2. Double-Shell Tank Liquid Means and 95% Upper Limits 4

Table 3-3. Pass/Fail Test for Double-Shell Tank Liquids 5

Table 3-4. Double-Shell Tank Solids Variability Ratios 7

Table 3-5. Double-Shell Tank Solids Means and 95% Upper Limits 7

Table 3-6. Pass/Fail Test for Double-Shell Tank Solids 8

LIST OF TERMS

DST	double-shell tank
SST	single-shell tank
TWINS	Tank Waste Information Network System
UL95	95% Upper Limit
%	percent
μCi/mL	microcuries per milliliter
μg/mL	micrograms per milliliter

1.0 INTRODUCTION AND BACKGROUND

The waste in 115 single-shell tanks (SSTs) and all 28 double-shell tanks (DSTs) has been "characterized" based on analytical data from core samples, grab samples, or auger samples. The analytical data are stored in the Tank Waste Information Network System (TWINS) database. TWINS contains approximately 1 million observations on radioactive isotopes and chemical elements. These analytical data are based on laboratory analysis of waste samples. Despite the abundance of data, 34 SSTs have not been sampled, and others have not been characterized to satisfy all current, or potential, characterization and disposal issues.

Because sampling and analysis of tank waste is an expensive enterprise, an incentive exists to reduce sampling and/or laboratory analytical work while at the same time obtaining needed data. Application of statistical techniques to the considerable volume of existing data offers some promise towards reducing sampling and/or analytical expenditures and personnel exposures.

This optimization study seeks to exploit the possibility of using statistical techniques by:

- Demonstration of the "proof of principle" of the statistical method to estimate concentration in unsampled tanks.
- Comparison of the "between tank" variability and the "within tank" variability.
- Estimation of the concentrations (and their uncertainty) of analytes, by liquid and solid waste phases, in unsampled tanks.
- Provision of an assessment of the viability of the method proposed.

The "proof of principle" is demonstrated on a subset of analytes from liquid and solid phases of DST waste. Cross-validation methods are used; i.e., omitting data from a tank and using the remainder to estimate the combined concentration means. Using statistical methods, the sample-based means and combined concentration mean estimates are compared.

Mean concentrations for many analytes have been estimated using data from TWINS. The uncertainty in the means is the "within tank" variability. The means and uncertainties are reported in the means and variances tables in TWINS. In this study, the existing means are statistically "combined" into a <u>single</u> mean that can be applied to all tanks in each phase. The uncertainty in this mean is the "between tank" variability.

Statistical methods are used to compare the "between tank" variability and the "within tank" variability. If they are approximately the same, then the uncertainty for the combined mean, which is applied to unsampled tanks, will be similar to that for tanks that have been sampled.

2.0 SUMMARY OF STATISTICAL TECHNIQUES

The means and variances tables in TWINS contain statistical summaries that have been used to analyze tank waste characterization data. These summaries contain an estimate of the mean analyte concentration, the standard deviation of the mean, and several variance components. The standard deviation of the mean is a combination of the within tank variability and the measurement variability. The existing summary statistics are classified by tank, by analyte, and by phase (solid or liquid). In this study, the means are statistically combined, by analyte, and by phase, into a single analyte mean that can be applied to all tanks. The uncertainty in this combined mean is a combination of the between tank variability and the within tank variability. Analysis of variance methods were used to estimate the combined mean analyte concentration and the standard deviation of the combined mean.

If the between tank variance is sufficiently small and if the combined and observed means are sufficiently close together, then the method is judged appropriate for estimating mean concentrations in the liquid or solid phases in DSTs. The definition of "sufficiently close" or "sufficiently small" depends on the requirements of the user of the data and invariably requires a cost trade off between accepting the variability of the data on the one-hand or obtaining additional data by sampling and laboratory analysis on the other.

Some values used in this study were below the detection limit. In such cases, the value of the detection limit was used. In addition, some data labeled "pre-transfer" in TWINS were used in this study.

3.0 STUDY RESULTS

3.1 DOUBLE-SHELL TANK LIQUID RESULTS

In Table 3-1, the between tank variance of standard chemicals of concern is divided by the within tank variance to arrive at a variance ratio. A ratio of 1 indicates the variances are the same. A ratio of less than one indicates the between tank variance is less than the within tank variance. A ratio of greater than one indicates the between tank variance is greater than the within tank variance.

These ratios are helpful in developing sampling strategies. For instance, for constituents near one where a reduction of uncertainty is needed by the project, a sampling strategy wherein it does not matter much what specific tank is sampled would be appropriate. In other words, tanks could be chosen that are relatively easy (less costly) to sample or wherein employee exposure can be minimized. It would not matter which specific tank(s) were chosen.

For constituents with a ratio well above one, a much more focused sampling strategy is required if uncertainty associated with the tank data is needed. In that case, the strategy might be to focus on sampling tanks with the greatest waste heterogeneity on the one hand and on those with the least heterogeneity on the other hand based on process knowledge. Such a strategy would result in a better understanding of waste variability both within tank and between tank.

Table 3-1. Double-Shell Tank Liquid Variability Ratios

Constituent	Variability Ratio	Constituent	Variability Ratio
Chemicals			
Aluminum	4.9	Nitrate	3.2
Bismuth	2.2	Nitrite	2.8
Calcium	4.3	Phosphate	3.4
Chloride	3.9	Potassium	9.1
Chromium	1.6	Silicon	1.1
Fluoride	3.8	Sodium	3.6
Iron	1.7	Sulfate	8.2
Lead	2.9	Uranium	1.2
Manganese	10.3	Zirconium*	0.4
Radionuclides			
Americium-241	1.1	Plutonium-239	5.1
Cesium-137	26.4	Selenium-79	5.7
Europium-154	5.6	Strontium-90	11.0
Iodine-129	3.1	Technetium-99	1.2

Table 3-1. Double-Shell Tank Liquid Variability Ratios

Constituent	Variability Ratio	Constituent	Variability Ratio
Neptunium-237	1.7		

Note: *Zirconium has only 21 percent detected values.

Table 3-2 provides the combined mean concentration value (one value fit for all tanks) for selected chemicals and radionuclides and the upper 95% limit (UL95) for the concentration values in μg/mL taken from TWINS sample data. This table is important for interpreting the following Table 3-3.

Table 3-2. Double-Shell Tank Liquid Means and 95% Upper Limits

Constituent	Mean (μg/mL)	UL 95 (μg/mL)
Chemicals		
Aluminum	1.32E+04	3.29E+04
Bismuth	4.27E+01	9.59E+01
Calcium	6.91E+01	2.75E+02
Chloride	3.15E+03	9.74E+03
Chromium	2.99E+02	8.10E+02
Fluoride	1.28E+03	6.40E+03
Iron	6.31E+01	6.01E+02
Lead	5.88E+01	1.98E+02
Manganese	1.80E+01	1.78E+02
Nitrate	8.42E+04	2.15E+05
Nitrite	5.50E+04	1.37E+05
Phosphate	2.22E+03	6.87E+03
Potassium	8.94E+03	3.24E+04
Silicon	1.29E+02	3.18E+02
Sodium	1.21E+05	2.78E+05
Sulfate	4.43E+03	1.52E+04
Uranium	1.80E+02	4.47E+02
Zirconium*	1.70E+01	9.29E+01
Radionuclides (μCi/mL)		
Americium-241	1.97E-01	9.28E-01
Cesium-137	3.34E+02	1.36E+03
Europium-154	7.87E-02	3.57E-01
Iodine-129	1.02E-04	2.89E-04
Neptunium-237	5.94E-01	2.64E+00
Plutonium-239	2.64E+03	6.16E+03
Selenium-79	4.45E-04	1.24E-03
Strontium-90	5.66E+00	4.43E+01
Technetium-99	1.02E-01	2.61E-01

*Note: Zirconium has only 21% detected values.

Table 3-3 provides an analysis of how often the best-basis inventory concentration value fell below the UL95 in Table 3-2 (passed) or was above the UL95 in Table 3-2 (failed).

Table 3-3. Pass/Fail Test for Double-Shell Tank Liquids

Constituent	Pass	Fail
Chemicals		
Aluminum	35	1
Bismuth	36	0
Calcium	34	2
Chloride	34	2
Chromium	34	2
Fluoride	33	3
Iron	35	1
Lead	35	1
Manganese	35	1
Nitrate	33	3
Nitrite	34	2
Phosphate	34	2
Potassium	35	1
Silicon	34	2
Sodium	36	0
Sulfate	34	2
Uranium	34	2
Zirconium*	35	1
Totals	620	28
Percentages	96%	4%
Radionuclides		
Americium-241	36	0
Cesium-137	34	2
Europium-154	35	1
Iodine-129	26	10
Neptunium-237	36	0
Plutonium-239	36	0
Selenium-79	33	3
Strontium-90	34	2
Technetium-99	30	6
Totals	300	24
Percentages	93%	7%

* Note: Zirconium has only 21% detected values.

The significance of Table 3-3 again comes down to the needs of each project. If the range between the mean and the UL95 in Table 3-2 is adequate and the pass rate found in Table 3-3 is

acceptable for the given project, additional sampling costs can be avoided for those constituents. If the project cannot stand the range or the pass rate, additional data may have to be collected from the tank. Cost of obtaining additional samples enters into the decision.

3.2 DOUBLE-SHELL TANK SOLIDS RESULTS

In Table 3-4, the between tank variance of standard chemicals of concern is divided by the within tank variance to arrive at a variance ratio. A ratio of 1 indicates the variances are the same. A ratio of less than one indicates the between tank variance is less than the within tank variance. A ratio of greater than one indicates the between tank variance is greater than the within tank variance.

These ratios are helpful in developing sampling strategies. For instance, for constituents near one where a reduction of uncertainty is needed by the project, a sampling strategy wherein it does not matter much what specific tank is sampled would be appropriate. In other words, tanks could be chosen that are relatively easy (less costly) to sample or wherein employee exposure can be minimized. It would not matter which specific tank(s) were chosen.

For constituents with a ratio well above one, a much more focused sampling strategy is required if uncertainty associated with the tank data is needed. In that case, the strategy might be to focus on sampling tanks with the greatest waste heterogeneity on the one hand and on those with the least heterogeneity on the other hand based on process knowledge. Such a strategy would result in a better understanding of waste variability both within tank and between tank.

Table 3-4. Double-Shell Tank Solids Variability Ratios

Constituent	Variability Ratio	Constituent	Variability Ratio
Chemicals			
Aluminum	1.1	Nitrate	1.6
Bismuth	0.3	Nitrite	2.1
Calcium	0.3	Phosphate	0.3
Chloride	5.2	Potassium	2.8
Chromium	0.5	Silicon	0.6
Fluoride	1.9	Sodium	2.7
Iron	8.6	Sulfate	0.5
Lead	1.1	Uranium	0.8
Manganese	2.7	Zirconium	3.4
Radionuclides			
Americium-241	1.7	Plutonium-239	2.7
Cesium-137	17.9	Selenium-79	1.2
Europium-154	2.1	Strontium-90	1.1
Iodine-129*	0.0	Technetium-99	1.3
Neptunium-237	1.7		

*Note: Iodine-129 has only 12% detected values.

Table 3-5 provides the combined mean concentration value (one value fit for all tanks) for selected chemicals and radionuclides and the upper 95% limit for the concentration values in μg/mL taken from TWINS sample data. This table is important for interpreting the following Table 3-6.

Table 3-5. Double-Shell Tank Solids Means and 95% Upper Limits

Constituent	Mean (μg/g)	UL 95 (μg/g)
Chemicals		
Aluminum	2.91E+04	5.86E+04
Bismuth*	1.01E+03	2.38E+03
Calcium	1.40E+03	3.17E+03
Chloride	3.92E+03	1.06E+04
Chromium	3.17E+03	9.83E+03
Fluoride	1.22E+04	5.59E+04
Iron	1.13E+04	6.27E+04
Lead	1.25E+03	3.08E+03
Manganese	1.76E+03	1.06E+04
Nitrate	9.19E+04	2.07E+05
Nitrite	6.62E+04	1.53E+05
Phosphate	6.18E+03	1.50E+04
Potassium	4.54E+03	2.35E+04
Silicon	1.36E+03	3.80E+03

Table 3-5. Double-Shell Tank Solids Means and 95% Upper Limits

Constituent	Mean (μg/g)	UL 95 (μg/g)
Sodium	1.69E+05	2.78E+05
Sulfate	7.72E+03	2.64E+04
Uranium	7.03E+03	2.24E+04
Zirconium	1.18E+04	7.31E+04
Radionuclides (μCi/g)		
Americium-241	1.50E+01	6.41E+01
Cesium-137	2.76E+02	9.84E+02
Europium-154	8.30E+00	5.20E+01
Iodine-129*	NA	NA
Neptunium-237	2.56E+01	1.42E+02
Plutonium-239	1.38E+00	4.82+00
Selenium-79	3.31E-03	1.72E-02
Strontium-90	2.47E+03	1.25E+04
Technetium-99	1.99E-01	2.81E-01

* Note: Iodine-129 has only 12% detected values.

Table 3-6 provides an analysis of how often the best-basis inventory concentration value fell below the UL95 in Table 3-5 (passed) or was above the UL95 in Table 3-5 (failed).

The significance of Table 3-6 again comes down to the needs of each project. If the range between the mean and the UL95 in Table 3-6 is adequate and the pass rate found in Table 3-5 is acceptable for the given project, additional sampling costs can be avoided for those constituents. If the project cannot stand the range or the pass rate, additional data may have to be collected from the tank. Cost of obtaining additional samples enters into the decision.

Table 3-6. Pass/Fail Test for Double-Shell Tank Solids

Constituent	Pass	Fail
Chemicals		
Aluminum	23	2
Bismuth	24	1
Calcium	22	3
Chloride	25	0
Chromium	21	4
Fluoride	23	2
Iron	22	3
Lead	24	1
Manganese	24	1
Nitrate	23	2
Nitrite	25	0
Phosphate	21	4
Potassium	24	1
Silicon	23	2

Table 3-6. Pass/Fail Test for Double-Shell Tank Solids

Constituent	Pass	Fail
Sodium	25	0
Sulfate	24	1
Uranium	23	2
Zirconium	23	2
Totals	419	31
Percentages	93%	7%
Radionuclides		
Americium-241	24	1
Cesium-137	24	1
Europium-154	25	0
Iodine-129*	NA	NA
Neptunium-237	25	0
Plutonium-239	24	1
Selenium-79	24	1
Strontium-90	24	1
Technetium-99	25	0
Totals	195	5
Percentages	97%	3%

*Note: I-129 has only 12% detected values.

4.0 POTENTIAL BENEFICIAL APPLICATIONS OF THE RESULTS

An examination of the results tables suggests a number of beneficial applications of this work including potential cost savings in sampling and laboratory analysis. The potential benefits are:

- Given the relatively high variability ratios for both chemicals and radionuclides in liquids, the sampling strategy for liquid grab samples should be targeted towards tanks with high concentrations of key constituents. In other words, a focused sampling strategy is suggested. The exception is Tc-99-- a very important radionuclide driving risk based cleanup activities. Because Tc-99 has a relatively low variability ratio, uncertainty in the data can be reduced using samples from any double-shell tank. In other words, additional Tc-99 data could be obtained opportunistically whenever any tank is sampled for any reason.

- For radionuclides in solids, Table 3-6 suggests that additional sampling may not be required for the key radionuclides listed if the projects can accept the range between means and the UL95 listed in Table 3-5. This could lead to avoidance of expensive core sampling for solids where the primary constituents of concern are radionuclides. In cases where cores are taken for other reasons, laboratory costs for analysis of Tc-99, Pu-239, and other long-lived radionuclides can be reduced.

5.0 RECOMMENDATIONS

Because of the promising results of this study particularly for radionuclides in solids, the following recommendations are made for additional work.

1. Extension of this phase 1 work to include SSTs and to include a longer list of chemicals and radionuclides.

2. Continuation of the work to phase 2 as outlined in the Optimization Study Task Plan. The additional phase 2 tasks include:

   - Evaluation of potential savings from Data Quality Objective (DQO) consolidation or elimination.
   - Estimation of potential savings from elimination of sampling of certain tanks.
   - Estimation of potential savings from reduction in application of laboratory analytical methods to certain analytes.
   - Determination of the overall cost reduction benefit from application of statistical methods.
   - Development of a regulatory strategy.
   - Presentation of regulatory strategy to the U.S. Department of Energy, Office of River Protection (ORP) and the regulators to gain approval for implementation.
   - Assessment of the operational risk associated with implementation of the developed strategy.

APPENDIX A

STATISTICAL METHODS EMPLOYED

Appendix A: Statistical Methods Employed

The model used to calculate the variance ratios is the one-way analysis of variance:

$$y_{ij} = \mu + b_i + e_{ij},$$

$$i = 1,2,\ldots a;\ j = 1,2,\ldots n_i$$

where "y_{ij}" is the jth laboratory measurement from double-shell tank "i," and μ represents the average for all double-shell tanks. It is assumed that the variables b_i and e_{ij} are uncorrelated with each other, have zero means, and variances σ_b^2, and σ_e^2. The variance, of observation "y", can be expressed as follows:

$$\sigma_y^2 = \sigma_b^2 + \sigma_e^2.$$

The ratio σ_b^2/σ_e^2 is a measure of the variance across tanks divided by the variance within tanks. When the ratio is less than one, it indicates that the majority of the variation is from differences between measurements within tanks. When the ratio is greater than one, most of the variation is due to changes in tank to tank concentration averages.

The S-PLUS™ software package (S-PLUS 2000) was used in the calculations. The variance ratios were calculated, in twelve steps, using the method given in Searle (1970), chapter 11. The steps, expressed in S-PLUS™ syntax, are as follows:

1. t0 <- sum(y^2)
2. yidot <- aggregate(y, by = b, FUN = sum)$x
3. ni <- aggregate(y, by = b, FUN = length)$x
4. ta <- sum(yidot^2/ni)
5. n <- length(y)
6. tu <- sum(y)^2/n
7. a <- length(b[!duplicated(b)])
8. s2 <- sum(ni^2)
9. s3 <- sum(ni^3)
10. s2e <- (t0 - ta)/(n - a)
11. s2b <- (ta - tu - (a - 1) * s2e)/(n - s2/n)
12. var.ratio <- s2b/s2e

where s2b and s2e represent, σ_b^2 and σ_e^2, respectively.

References:

Searle, S. R., 1970, *Linear Models*, Wiley, New York.

S-PLUS, 2000, *S-PLUS 2000*, Data Analysis Product Division, MathSoft, Inc., Seattle, Washington.