DOCUMENT RELEASE FORM

(1) Document Number: RPP-6924	(2) Revision Number: 1	(3) Effective Date: June 1, 2010	
(4) Document Type:	☐ Digital Image ☐ Hard copy ☒ PDF ☐ Video	(a) Number of pages (including the DRF) or number of digital images	14
(5) Release Type	☐ New ☐ Cancel	☐ Page Change	☒ Complete Revision
(6) Document Title:	Statistical Methods for Estimating the Uncertainty in the Best Basis Inventories		
(7) Change/Release Description:	Consolidate description of statistical methods currently being used for BBI uncertainties into one document. Remove obsolete information.		
(8) Change Justification:	Simplify documentation. Ensure document accuracy.		

(9) Associated Structure, System, and Component (SSC) and Building Number:	(a) Structure Location:	(c) Building Number:	(e) Project Number:
	N/A	N/A	N/A
	(b) System Designator:	(d) Equipment ID Number (EIN):	
	N/A	N/A	

(10) Impacted Documents:	(a) Document Type	(b) Document Number	(c) Document Revision
	N/A		

(11) Approvals:

(a) Author (Print/Sign): D. A. Greer [signature]	Date: June 1, 2010
(b) Responsible Manager (Print/Sign): J.G. Reynolds [signature]	Date: June 1, 2010
(c) Reviewer (Optional, Print/Sign): D.M.Nguyen [signature]	Date: June 1, 2010
(d) Reviewer (Optional, Print/Sign):	Date:

(12) Distribution:

(a) Name	(b) MSIN	(a) Name	(b) MSIN	Release Stamp
M.A. Anderson	B1-55	J.G. Reynolds	B1-55	JUN 0 1 2010
D.L. Banning	B1-55	M.J. Rodgers	B1-55	DATE:
R.S. Disselkamp	B1-55	N.W. Kirch	R2-58	HANFORD RELEASE
S.J. Harrington	B1-55	P.L. Rutland	B1-55	STA: 15
D.M. Nguyen	B1-55	R.J. Koll	H6-60	ID: 65
D.E. Place	B1-55			
J.H. Rasmussen	B1-55			

(13) Clearance	(a) Cleared for Public Release	(b) Restricted Information?	(c) Restriction Type:
	☒ Yes ☐ No	☐ Yes ☒ No	

(14) Clearance Review (Print/Sign):	Date:
NANCY A FOUAD / Nancy A Fouad	6-1-10

Statistical Methods for Estimating the Uncertainty in the Best Basis Inventories

D.A. Greer
Washington River Protection Solutions LLC
Richland, WA 99352
U.S. Department of Energy Contract DE-AC27-08RV14800

EDT/ECN: DRF UC: None
Cost Center: 2GB00 Charge Code:
B&R Code: None Total Pages: 14

Key Words: statistical method, uncertainty, best basis inventory, BBI, analyte concentration, waste density, waste volume, confidence interval, relative standard deviation, RSD, Searle equation

Abstract: This document describes the statistical methods used to determine uncertainty estimates for the Best Basis Inventory (BBI). For each waste phase, the inventory is calculated as the product of the concentration, the density, and the volume. The squared relative uncertainty in the waste phase inventory is the sum of the squared relative uncertainty in each of the concentration, density and volume for that phase.The relative uncertainty in the volume is based on the method used to measure the waste height, hence volume. The relative uncertainty of the density is based on the phase of the material (solid or liquid) and the type of tank (single or double shell). The relative uncertainty of the concentration is calculated using the Searle equations on tank sample data.The total inventory is the sum of the waste phase inventories. The uncertainty in the total inventory is the sum of the uncertainty in each waste phase.

Nancy A Fouad 6-1-2010
Release Approval Date

JUN 01 2010
DATE:
STA: 15
HANFORD RELEASE
ID: 65
Release Stamp

Approved For Public Release

Tank Operations Contractor (TOC) **RECORD OF REVISION**	(1) Document Number: **RPP-6924**	Page 1

(2) Title:

Statistical Methods for Estimating the Uncertainty in the Best Basis Inventories

Change Control Record

(3) Revision	(4) Description of Change – Replace, Add, and Delete Pages	Authorized for Release	
		(5) Resp. Engr. (print/sign/date)	(6) Resp. Mgr. (print/sign/date)
0	Initial Release – EDT 630479	L. Jensen September 7, 2000	J.G. Field September 7, 2000
1 RS	Complete revision with DRF.	D.A. Greer June 1, 2010	J.G. Reynolds June 1, 2010

Statistical Methods for Estimating the Uncertainty in the Best Basis Inventories

D.A. Greer
Washington River Protection Solutions LLC

Date Published
June 2010

Post Office Box 850
Richland, Washington

Prepared for the U.S. Department of Energy
Office of River Protection

Contents

Appendix

List of Terms

Abbreviations

C	Concentration
Cov	Covariance
D	Density
I	Inventory
RSD	Relative Standard Deviation
SD	Standard Deviation
V	Volume

Units

μCi/mL	microcuries per milliliter
μg/mL	micrograms per milliliter

1.0 INTRODUCTION

This document describes the methods used to provide estimates of the uncertainty of the inventory amounts in the Best Basis Inventory. The uncertainty value that is provided will be used to create an interval, similar to a statistical confidence interval, for the estimated inventory value. However, due to the construction of the uncertainty by varying methods, some of which do not include sample data, this interval is not a confidence interval. In general, the interval is constructed as the inventory value plus or minus two times the standard deviation (SD) of the inventory value. This interval should be interpreted as a general guideline reflecting the uncertainty in the inventory value, but there is no specific probabilistic value associated with it.

2.0 WASTE PHASE INVENTORY UNCERTAINTY

Tank waste is divided into three phases: sludge, saltcake and supernatant. For each phase, the equation for the inventory of an analyte is given by:

$$I = C \times D \times V$$

where I represents the inventory, C represents the analyte concentration, D represents the waste density, and V represents the volume of the waste in the particular phase. Note that the term "density" is used to denote "Bulk Density" or "Specific Gravity", depending on whether the waste phase is solid or liquid.

Assuming that the three variables C, D and V are independent of each other, the uncertainty in the inventory is calculated as follows:

$$RSD_I^2 \cong RSD_C^2 + RSD_D^2 + RSD_V^2$$

where RSD is defined as the relative standard deviation (RSD), i.e., the SD divided by the mean. The subscript identifies the specific variable with which the RSD is associated. Also, for liquids, the density is not used in the inventory calculation, since the concentrations are reported on a volumetric basis (µg/mL or µCi/mL). Thus, the corresponding RSD is omitted from the uncertainty calculation.

See Appendix 1 for the derivation of the uncertainty of the waste phase inventory.

3.0 TOTAL INVENTORY UNCERTAINTY

Let I_{sludge}, $I_{saltcake}$, and $I_{supernatant}$ denote the inventory in the sludge, saltcake and supernatant phases, respectively, for a particular analyte. Then, total inventory is calculated as:

$$I_{Total} = I_{Sludge} + I_{Saltcake} + I_{Supernatant}$$

Let SD(I_{Sludge}), SD($I_{Saltcake}$), and SD($I_{Supernatant}$) denote the SD of the inventory estimates for the sludge, saltcake and supernatant phases, respectively. Then the uncertainty in the total inventory is calculated as:

$$SD(I_{Total}) = SD(I_{Sludge}) + SD(I_{Saltcake}) + SD(I_{Supernatant})$$

See Appendix 2 for the derivation of SD(I_{Total}).

4.0 VOLUME UNCERTAINTY

The uncertainty in the waste volume in a tank is based on the uncertainty in the height of each phase of waste material in the tank. For the waste phases of sludge, saltcake and supernatant, there are seven different combinations which can be present in a tank. The following table lists the seven cases, the associated uncertainty in the waste phase height measurement (representing one SD in inches), and the measurement method used as a source for the uncertainty. Note that these SDs are applicable to the uncertainty in a waste phase, not the uncertainty in the total volume of waste in a tank. See Jensen and Wilmarth, Volume Uncertainties for Best Basis Inventories, 2002 for further explanation of the analysis producing the table.

One Standard Deviation (SD) in Inches of Waste Phase Height Measurement				
Case	**Sludge**	**Saltcake**	**Supernatant**	**Basis for Uncertainty**
1	—	—	0.25	ENRAF
2	4.3	—	—	Sludge Tanks
3	6.5	—	6.5	Solids Level Measurement
4	—	11.5	—	Saltcake Tanks
5	—	6.5	6.5	Solids Level Measurement
6	9.5	9.5	—	Half Segment
7	9.5	9.5	6.5	Half Segment and Solids Level Measurement

In order to use these uncertainties, they first need to be converted from height to volume. For a 75 foot diameter tank, one inch in height is approximately equal to 2,750 gallons in volume. Therefore, one SD in volume uncertainty for a 75 foot diameter tank is obtained by multiplying the table value by 2,750. To obtain a volume RSD, the volume SD is divided by the volume.

In some cases, other procedures are used to provide an estimate of the waste volume. The uncertainty associated with the specific procedure should be used to estimate the volume uncertainty, rather than the table value.

5.0 DENSITY UNCERTAINTY

If multiple density values are available from tank sample data, then the mean and SD of the sample data are calculated using the Searle equations described below.

In many cases, sample data are not available for estimating the uncertainty of density. In this situation, the following table is used to provide estimates of the density uncertainty, based on previous analysis. See Appendix 3 for more details on the analysis which produced the table.

Relative Standard Deviations (RSDs) for Bulk Density and Specific Gravity by Type of Tank and Waste Phase		
Tank and Phase	**Method**	**RSD (%)**
SST, Liquid	Specific Gravity	5.90%
SST, Solid	Bulk Density	7.55%
DST, Liquid	Specific Gravity	8.16%
DST, Solid	Bulk Density	6.50%

6.0 CONCENTRATION UNCERTAINTY

The following equations (referred to as the "Searle Equations", from Searle, Casella, and McCulloch, 1992, page 428) provide a closed-form solution to the unbalanced one-way analysis of variance situation for estimating the "treatment" (lab sample id) and "error" (primary/duplicate/triplicate) variance components. These variance component estimates are then incorporated into weights used in the Generalized Least Squares Estimator methodology (Searle, Casella, and McCulloch, 1992, page 54) to produce the estimated overall mean and SD.

Calculate

$$N = \sum_{i=1}^{a} n_i$$

$$S_2 = \sum_{i=1}^{a} n_i^2$$

$$T_O = \sum_{i=1}^{a} \sum_{j=1}^{n_i} y_{ij}^2$$

$$T_A = \sum_{i=1}^{a} \frac{y_{i.}^2}{n_i}$$

$$T_\mu = \frac{y_{..}^2}{N}$$

where N is the total number of measurements, n_i is the number of measurements in each level, and y_{ij} represents the actual measurement, with subscript i indicating the lab sample id level, subscript j indicating the primary, duplicate, or triplicate level of measurement. A dot (•) in the subscript indicates a sum over the range of that subscript.

Then the variance associated with the primary/duplicate/triplicate measurements is calculated as:

$$\hat{\sigma}_e^2 = \frac{(T_O - T_A)}{(N - a)}$$

The variance associated with the lab sample id is calculated as:

$$\hat{\sigma}_\alpha^2 = \frac{\left(T_A - T_\mu - (a-1) \times \hat{\sigma}_e^2\right)}{\left(N - \frac{S_2}{N}\right)}$$

Then, calculate the weights (w_i) and the averages ($\bar{y}_i$) at the lab sample id level:

$$w_i = \frac{1}{\hat{\sigma}_\alpha^2 + \frac{\hat{\sigma}_e^2}{n_i}}$$

$$\bar{y}_i = \sum_{j=1}^{n_i} \frac{y_{ij}}{n_i}$$

Finally, the overall average and SD are calculated as:

$$average = \frac{\sum_{i=1}^{a} w_i \times \bar{y}_i}{\sum_{i=1}^{a} w_i}$$

$$SD = \sqrt{\frac{1}{\sum_{i=1}^{a} w_i}}$$

In some cases, observations for a given analyte are below the detection limit. If 50% or more are above the detection limit, then the detection limit is used as a quantitative value in computing the mean and variance. If less than 50% are above the detection limit, then no uncertainty estimates are computed. In this case a default RSD of 100% is used. (A justification for using 100% as the default value of the RSD is as follows. Let the random variable X be defined on the interval [0, 1] and let its mass be concentrated at 0 and 1 with probabilities p and 1-p; i.e., “0” represents below the detection limit and “1” above the detection limit. The mean of X is p and the variance is p*(1-p). The maximum value of the variance occurs when p=0.5. Consequently the variance is less than $(0.5)^2$ and the RSD is less than 100%.) Also, if there is only one observation a default RSD of 100% is used.

7.0 REFERENCES

Hogg, R. V., and A. T. Craig, 1971, *Introduction to Mathematical Statistics*, third edition, The Macmillan Company, New York, New York.

Jensen, L., and S.R. Wilmarth, Volume Uncertainties for Best-Basis Inventories, April 5, 2002, Interoffice Memo 7G300-02-JGF-003

Searle, S. R., G. Casella, and C. E. McCulloch, 1992, *Variance Components*, third edition, John Wiley & Sons, Inc., New York, New York.

APPENDIX 1 – DERIVATION OF UNCERTAINTY OF WASTE PHASE INVENTORY

Tank waste is divided into three phases: sludge, saltcake and supernatant. For each phase, the equation for the inventory of an analyte is given by:

$$I = C \times D \times V$$

where I represents the inventory, C represents the analyte concentration, D represents the waste density, and V represents the volume of the waste in the particular phase.

Assuming that the three variables C, D and V are independent of each other, the uncertainty in the inventory is calculated as follows (Hogg and Craig 1971, page 168):

$$RSD_I^2 = RSD_C^2 + RSD_D^2 + RSD_V^2 + RSD_C^2 \times RSD_D^2 + RSD_C^2 \times RSD_V^2 + RSD_D^2 \times RSD_V^2 + RSD_C^2 \times RSD_D^2 \times RSD_V^2$$

where RSD is defined as the relative standard deviation, i.e., the standard deviation (SD) divided by the mean. The subscript identifies the specific variable with which the RSD is associated.

Ignoring the higher-order product terms in the equation produces the following approximation:

$$RSD_I^2 \cong RSD_C^2 + RSD_D^2 + RSD_V^2$$

APPENDIX 2 – DERIVATION OF UNCERTAINTY OF TOTAL INVENTORY

For the more general solution, let I_1, I_2, …, I_k be the inventory estimates for waste phases 1, 2, …, k, respectively. Additionally, let $SD(I_1)$, $SD(I_2)$, …, $SD(I_k)$ be the standard deviations of the inventory estimates for waste phases 1, 2, …, k, respectively. Then, the total inventory is calculated using the following:

$$I_{Total} = \sum_{i=1}^{k} w_i \times I_i$$

where the w_i are known "weights" (which are all equal to one, unless otherwise noted). Then the variance of the total inventory (where variance is the square of the standard deviation) is:

$$SD^2(I_{Total}) = \sum_{i=1}^{k} w_i^2 SD^2(I_i) + 2 \times \sum_{i=1}^{k} \sum_{j=i+1}^{k} w_i \times w_j \times cov(I_i, I_j)$$

where $cov(I_i,I_j)$ represents the covariance between the inventories in the i^{th} and j^{th} waste phases. In general, these covariances are unknown. However, since the correlation between two variables is always less than or equal to one in absolute value (Hogg and Craig 1971, page 75), the covariance will be less than or equal to the product of the two standard deviations, i.e.,:

$$cov(I_i, I_j) \leq SD(I_i) \times SD(I_j)$$

Using this, an upper bound on the variance of the total inventory can be calculated as:

$$SD^2(I_{Total}) \leq \sum_{i=1}^{k} w_i^2 \times SD^2(I_i) + 2 \times \sum_{i=1}^{k} \sum_{j=i+1}^{k} w_i \times w_j \times SD(I_i) \times SD(I_j)$$

This can be further simplified algebraically, as follows:

$$SD^2(I_{Total}) \leq \left\{ \sum_{i=1}^{k} w_i \times SD(I_i) \right\}^2$$

This then results in:

$$SD(I_{Total}) \leq \sum_{i=1}^{k} w_i \times SD(I_i)$$

Letting i represent sludge, saltcake and supernatant, and using all w_i=1, the following approximation is used for the uncertainty of the total inventory:

$$SD(I_{Total}) \cong SD(I_{Sludge}) + SD(I_{Saltcake}) + SD(I_{Supernatant})$$

APPENDIX 3 – SUMMARY OF ANALYSIS TO ESTIMATE DENSITY UNCERTAINTY

Using data from the Tank Characterization Database (TCD), estimates of the mean density D and standard deviation of the mean (SD(D)) were obtained for each type of tank, single-shell tank (SST) or double-shell tank (DST), and waste phase, liquid or solid. All of the bulk density and specific gravity observations available in 2000 in TCD were obtained. The observations were separated into four classes by type of tank (SST or DST) and waste phase (liquid or solid). For each of the four classes, a one-way analysis of variance model was fit to the data. The classification variable was the tank name. The results of the analysis of variance are given in the following table. Specifically, the analysis of variance was used to estimate the "pooled" within tank variance. This variance, often called the "error variance," is the S^2 (error) column in the table. In this table, the column Number of Tanks reports the number of tanks with Bulk Density or Specific Gravity observations that were used in the analysis, and the Mean of Means column is the un-weighted average of the Bulk Density or Specific Gravity tank means. The RSD (%) is the square root of S^2 (error), divided by the Mean of Means, times 100. These RSDs are used for each of the four tank/waste phase classifications.

Relative Standard Deviations (RSDs) for Bulk Density and Specific Gravity by Type of Tank and Waste Phase

Tank and Phase	Method	Number of Tanks	Mean of Means	S^2 (error)	RSD (%)
SST, Liquid	Specific Gravity	52	1.3293	0.006146	5.90%
SST, Solid	Bulk Density	60	1.5682	0.014001	7.55%
DST, Liquid	Specific Gravity	26	1.2036	0.009642	8.16%
DST, Solid	Bulk Density	18	1.5047	0.009553	6.50%