

Hanford Single-Shell Tank Grouping Study

KM Remund
CM Anderson
Pacific Northwest Laboratory

BC Simpson
Westinghouse Hanford Company

Prepared for Westinghouse Hanford Company

October 1995

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

al

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. 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. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

**Portions of this document may be illegible
electronic image products. Images are
produced from the best available original
document.**

Executive Summary

A tank grouping study has been conducted to find Hanford single-shell tanks with similar waste properties. The limited sampling resources of the characterization program could be allocated more effectively by having a better understanding of the groups of tanks that have similar waste types. If meaningful groups of tanks can be identified, tank sampling requirements may be reduced, and the uncertainty of the characterization estimates may be narrowed.

This tank grouping study considers the analytical sampling information and the historical information that is available for all single-shell tanks. The two primary sources of historical characterization estimates and information come from the Historical Tank Content Estimate (HTCE) Model and the Sort on Radioactive Waste Tanks (SORWT) Model. The sampling and historical information are used together to come up with meaningful groups of similar tanks.

Based on the results of analyses presented in this report, credible tank grouping looks very promising. Some groups defined using historical information (HTCE and SORWT) correspond well with those based on analytical data alone.

An overview of the results from the FY95 efforts follows:

1. Based on the analytical data, it was observed that tanks containing 1C and 2C44-51 (see list of terms) wastes generally grouped together. This supports suggestions put forth by LANL and WHC scientists that these two waste types are similar, despite the fact that the historical models generate a distinction between the two.
2. The available analytical data was used to perform an evaluation of the tank groups based on the two process history models. At present, there is not enough analytical information to perform a complete evaluation of the historical groupings. However, the historical groupings were consistent with some general grouping patterns in the analytical data. For example, when the analytical data showed clear separation between some salt cake and sludge tanks, historical groupings showed the same separation patterns.
3. An investigation was conducted to determine if the list of Hanford Defined Wastes could be reduced from 50 to some smaller number (one suggestion was that there were really only 8 to 10 distinct Hanford waste types). The results of the cluster analysis indicate that 23 of the 50 waste types were distinct. The remaining 27 waste types grouped together as 12 pairs and a triplet, or 13 groups. For example, saltcakes generated from B campaign and T1 campaign were similar. This result may allow for a small reduction (i.e., from 50 waste types to 36 distinct waste types) in the number of Hanford Defined Wastes, but not the large reduction that has been suggested.
4. The SORWT and HTCE Models were compared by performing a cluster analysis on the HTCE estimates and then comparing the tank groups based on the two models. There were some reasonable similarities between the models on a tank-by-tank comparison. On a group-by-group level, the comparison was not as reasonable. Only 7 of the 24 (25%) SORWT groups¹ remained intact after the HTCE cluster analysis was run. There were only 37% of the single shell tanks that grouped similarly based on the two models. However, using Figure 1 other intersections between the two grouping methods become apparent (i.e., SORWT Group 1 and HTCE Clusters 4 and 20). This information can be used to guide future sampling efforts and data interpretation.
5. A demonstration is presented showing how combined estimates (i.e., combined across tanks) of spatial variability can be calculated. As an example, pooled spatial variability estimates were calculated combining analytical results from Tanks B-110, B-111, BX-107 and T-104. These tanks were selected because they grouped together based on TCR mean concentrations and spatial variabilities.

This report should be viewed as a demonstration of how a more complete tank grouping study should be conducted when the analytical data becomes available. This is due to the fact that only 15 of the tanks

¹ There are 40 SORWT groups but only 24 of them include more than one tank.

considered in FY95 had core sampling results. The remaining tanks included in this years efforts used Braun data which is a less reliable source than the core sampling data. This effort will continue to be funded in FY96. It is expected that approximately 25 additional core sampled tanks will be added to the study in FY96. With the addition of these tanks and others in the years to come, more well defined tank grouping patterns will be identified. These tank grouping patterns can help determine how future sampling resources can be allocated more efficiently.

An updated comparison between the HTCE Model estimates and analytical results will be possible in FY96 as well. This is due to the refinements that Los Alamos National Laboratory is making to the HTCE Model.

Contents

1	Overview	1
1.1	Introduction	1
1.2	Scope	2
2	Process History Grouping Models	3
2.1	HTCE and SORWT Comparisons	3
2.2	Hanford Defined Wastes Cluster Analysis	10
3	Tank Grouping Efforts Based on Analytical Data	13
3.1	Analytical Data Description	13
3.2	Tank Grouping Analysis Based on TCR Mean Concentration Estimates	15
3.2.1	TCR Cluster Analysis Using Normalized Means	15
3.2.2	TCR Cluster Analysis Using Raw Means	16
3.3	Tank Grouping Analysis Based on TCR Spatial Variability Estimates	17
3.4	Tank Grouping Analysis Based on Mean Concentrations of All Analytical Results	18
3.4.1	Global Cluster Analysis on Normalized Means	19
3.4.2	Global Cluster Analysis on Raw Means	22
4	Between and Within Cluster Similarities	25
5	Pooled Estimates of Spatial Variability	29
6	Conclusions and Suggestions for Further Work	31
7	References	33
A	Data Description	35
A.1	TCR Estimates	35
A.2	TWINS Data	35
A.3	Braun Data	35
A.4	Missing Data	36
A.5	Data Plots	38

List of Figures

1	Summary of HTCE Cluster Analysis	9
2	Tank Grouping Data Collection	14
3	Illustration of Tank Means for Aluminum and Bismuth	39
4	Illustration of Tank Means for Iron and Sodium	40
5	Illustration of Tank Means for Cesium-137 and Strontium-90	41
6	Illustration of Tank Means for Uranium and Silicon	42
7	Illustration of Tank Means for Phosphate and Manganese	43
8	Illustration of Tank Means for Chromium and Nitrite	44
9	Illustration of Tank Means for Nitrate and Fluoride	45
10	Illustration of Tank Means for Water	46

List of Tables

1	Tanks With Identical Estimates from HTCE Model	4
2	Selected Analytes for Tank Grouping Study	4
3	Cluster Analysis Results for HTCE Data	5-8
4	Hanford Defined Wastes Cluster Analysis Results	10
5	Cluster Analysis Results for Normalized TCR Mean Concentration Estimates	16
6	Cluster Analysis Results for TCR Mean Concentration Estimates (Constituents Weighted According to Magnitude)	17
7	Cluster Analysis Results for the TCR Relative Standard Deviations	18
8	Cluster Analysis Results for Normalized Global Analytical Dataset	20-22
9	Cluster Analysis Results for Proportionally Weighted Global Analytical Dataset	22-24
10	ANOVA Results for 14 Tanks in 10 Groups (Across Groups)	26
11	ANOVA Results for 1C/2C Waste Type Group (Within Group)	26
12	Cluster Analysis Results from Section 3	29
13	Pooled and Tank Spatial Variability Comparisons	30
14	Braun Analytical Results Extreme Value Cutoff Points	36
15	Division of Braun Sample Types	36

List of Terms

ANOVA	Analysis of variance
CH statistic	Calinski-Harabaz Statistic
EM Algorithm	Expectation Maximization Algorithm
HDW/TLM	Hanford Defined Wastes/Tank Layering Model
HTCE	Historical Tank Content Estimate
ICP	Inductively coupled plasma atomic emission spectrometry
LANL	Los Alamos National Labs
MANOVA	Multivariate analysis of variance
PCA	Principal Components Analysis
RSD	Relative Standard Deviation
SORWT	Sort on Radioactive Waste Tanks
TCD	Tank Characterization Database
TCR	Tank Characterization Report
TRAC	Track of Radioactive Waste Model
TWINS	Tank Waste Information Network System
WHC	Westinghouse Hanford Company

SORWT Model Waste Types

1C	first cycle decontamination waste
1C-F	first cycle decontamination waste with ferrocyanide-scavenged waste
224	lanthanum fluoride decontamination waste
2C	second cycle decontamination waste
5-6	high-level B Plant waste from the bottom of Section 5
CCPLX	complex concentrate
CW	cladding waste
DIA	diatomaceous earth
DSSF	double-shell slurry feed
DW	decontamination waste
EB	evaporator bottoms
EB-ITS	in-tank solidification evaporator bottoms
EVAP	post-1976 evaporator feed
F	ferrocyanide-scavenged waste
HS	hot semiworks waste
ITS	in-tank solidification
IX	ion exchange waste
MIX	mixture of several miscellaneous wastes
NCPLX	noncomplexed waste
OWW	organic wash waste
P	neutralized acid waste
R	high-level REDOX waste
REDOX	reduction/oxidation
SLUICE	water remaining from sluicing activity
SRS	strontium leached sludge
SR-WASH	particulates from strontium wash of PUREX wastes in the AR vault
TBP	tributyl phosphate
TBP-F	tributyl phosphate ferrocyanide scavenged waste
UK	unknown

HTCE Model Waste Types

1C44-51/CW	1st cycle decontamination from bismuth phosphate process, 1944 through 1951 - includes cladding waste
1C52-56/CW	1st cycle decontamination from bismuth phosphate process, 1952 through 1956 - includes cladding waste
1CFeCN	ferrocyanide scavenged, 1st decontamination process supernates
224	lanthanum fluoride finishing waste from 224-U
2C44-51	2nd cycle decontamination waste from bismuth phosphate process, 1944 through 1951
2C52-56	2nd cycle decontamination waste from bismuth phosphate process, 1952 through 1956 - includes supernates formerly cribbed at T Plant
A1-SC	salt cake from evaporator campaign A1
A2-SS	salt slurry from evaporator campaign A2
AR	AR Solids = washed Purex sludge - also used to derive strontium recovery waste
B	high level waste from Purex Acidified Waste (PAW) - also refers to aluminum clad fuel, processed through B Plant for strontium extraction
BL	low level waste from all operations
B-SC	salt cake from evaporator campaign B
BY-SC	salt cake from evaporator campaign BY
CC1	complexant concentrate
CSR	cesium recovery waste
CW/Zr66-72	zirconium cladding wastes, 1966 to 1972
CWP/Al56-60	aluminum cladding Purex wastes, 1956 to 1960
CWP/Al61-72	aluminum cladding Purex wastes, 1961 to 1972
CWP/Zr83-88	zirconium cladding wastes, 1983 to 1988
CWR50-61	Redox cladding wastes, 1950 to 1961
CWR/Al52-60	aluminum cladding Redox wastes, 1952 to 1960
CWR/Al61-67	aluminum cladding Redox wastes, 1961 to 1967
DC	dilute complex waste
DW	decontamination waste
MW44-51	metal waste from bismuth phosphate process, 1955 through 1951
MW52-56	metal waste from bismuth phosphate process, 1952 through 1956
N	N-Reactor wastes
OWW55-65	Purex organic wash waste, 1955 to 1965
OWW69-72	Purex organic wash waste, 1969 to 1972
P56-62	Purex high level waste, 1956 through 1962
P63-67	Purex high level waste, 1963 through 1967
PFeCN1	ferrocyanide sludge produced by in-plant scavenging of waste from uranium recovery, using 0.005 M ferrocyanide
PFeCN2	same as PFeCN1, except using 0.0025 M ferrocyanide
PL	Purex low-level waste
PL83-88	Purex low-level waste, 1983 to 1988

HTCE Model Waste Types, Continued

R52-58	redox waste, 1952 through 1958
R59-67	redox waste, 1959 through 1967
R-SC	salt cake from evaporator run of redox wastes
S1-SC	salt cake from evaporator campaign S1
S2-SS	salt slurry from evaporator campaign S2
SRR	strontium recovery waste
SRR/FMJ	strontium recovery waste from sluiced Purex sludge
T1-SC	salt cake from evaporator campaign T1
T2-SC	salt cake from evaporator campaign T2
TFeCN	ferrocyanide sludge produced by in-tank or in-farm scavenging in C
TH66	1966 thoria campaign waste
TH70	1970 thoria campaign waste
UR/TBP	tri-butyl phosphate waste from solvent-based uranium recovery op- eration in 1950s
Z/PFP	234-5Z waste from Z-Plant plutonium finishing

Acknowledgements

We acknowledge the contribution that the Westinghouse Hanford Company statisticians made in this effort. They provided much of the analytical data that was used to obtain pooled estimates of spatial variability.

We also acknowledge the assistance that we received from Steve Agnew and his staff at Los Alamos National Laboratories in obtaining the Historical Tank Content Estimates for our grouping studies.

1 Overview

1.1 Introduction

The single-shell tanks at the Hanford site contain waste from special nuclear material production activities conducted at the site since 1944. The characterization of this waste is vital to future efforts to retrieve, pretreat, and dispose of the waste in an acceptable and safe manner. Tank waste characterization has been difficult due to the hazardous and complex nature of the waste.

There have been two independent approaches to characterizing the waste in these tanks. The first approach involved taking samples of the waste by any of several sampling methods (e.g., core sampling, auger sampling), and analyzing these samples for various chemical, physical and radiological constituents. The analytical results from these tank waste samples are used to estimate the waste characteristics of the whole tank. Several issues indicate that the samples taken from the tanks may not be typical of the tank waste as a whole. These issues include:

1. Limited accessibility to sample the waste. Usually only a few riser locations are available for sample access.
2. The sampler methods (i.e., push-mode, rotary) may not achieve complete recovery and thus may create biases in the data.
3. Depending on past tank transfer activities, the waste that lies below the risers could be atypical of the waste in the rest of the tank.
4. Small sample sizes in relation to the size of a tank (2 one-inch diameter core samples from a tank 75-feet in diameter) may not be adequate.
5. The waste within the tank may be heterogeneous (horizontally, as well as vertically).

Consequently, many have questioned the validity of the characterization estimates from this sampling approach.

This sampling-based approach to tank waste characterization will consume much time and resources if every tank must be sampled.

The second approach to tank waste characterization is based on the large amount of historical information that is available regarding the waste stored in the tanks. This historical information includes process histories, tank waste management records, and other data. This historical information is coupled with chemical knowledge of how wastes interact, how they age, and how they react with other added materials. The basic approach is to determine what wastes were put into a tank and how those wastes were treated and moved around to other tanks. As with the sampling characterization approach, there are limitations associated with this process history-based characterization approach. Although a vast amount of waste process historical information is available, it is still incomplete. Some waste transfers between tanks were not recorded. Also, some physical assumptions had to be made in the models which oversimplify some of the complex chemical reactions in the waste.

This tank grouping study was conducted to determine if tanks could be identified that have similar solid waste properties². These groups of similar tanks could improve the tank characterization effort in two ways. First, identifying similar tanks will help facilitate multivariate (across analyte) comparisons between the sampling and history based characterization approaches. Second, the limited resources that exist for sampling efforts can be used more efficiently. For example, if a group of similar tanks is identified, a small number of those tanks may only need to be sampled to make inferences to the whole group. This will save the sampling resources that would have been consumed by sampling the remaining tanks in the group.

²This tank grouping study is confined to analysis of solid waste samples since they are of primary interest in the tank characterization effort.

1.2 Scope

Three main objectives for this effort in FY95 are as follows:

1. To compare the two different process history models. Each process history model explicitly or implicitly groups tanks based on the waste types the tanks received. This part of the effort will identify instances when the grouping patterns for the different models are consistent or inconsistent.
2. To evaluate tank groupings based on the process history models, using the available analytical data. The analytical data from Hanford tank waste samples can also be used to define groups of tanks with similar waste characteristics. These tank groups are independent of the process history models. The sampling-based tank groupings can thus be used to evaluate tank groupings based on the process history models.
3. To use the sampling results from the grouped tanks to form pooled estimates of spatial variability that are general to a group of tanks as a whole. These pooled estimates of spatial variability would be based on more sampling information (more core samples) than the spatial variability estimates based on a few samples (typically two) from a single tank.

This tank grouping effort was confined to the single-shell tanks at Hanford for FY95, since these tanks are at present the focus of the process history models.

This report should be viewed as a demonstration of how a more complete tank grouping study should be conducted when the analytical data becomes available.

2 Process History Grouping Models

Several models have been used to determine the waste characteristics of the tanks, based on historical information. These include the Track Radioactive Wastes Model (TRAC) [7], the Sort On Radioactive Waste Tanks (SORWT) Model [6], and the Hanford Defined Wastes/Tank Layering Model (HDW/TLM) [1]. The HDW/TLM model has been used to obtain Historical Tank Content Estimates (HTCE) [3]. For simplicity, the HDW/TLM model will be referred to as the HTCE Model.

The TRAC Model was the first process history model that provided waste inventory estimates for the 149 single-shell tanks. Past comparisons with TRAC estimates and sampling results showed a large measure of disagreement. The SORWT and HTCE Models have replaced TRAC as the process history models that are being used in the characterization effort at present.

The SORWT model presents a method to categorize Hanford single shell tanks based on the major waste types that were deposited in them, and on the process histories of each tank. Tanks were characterized by noting the volume percentages of each waste type in each tank. The waste type that comprised the largest volume in a tank was designated as the primary waste type for that tank. Waste types comprising the second and third largest volumes were designated as that tank's secondary and tertiary types, respectively. Tanks that had the same primary and secondary waste types were then grouped together. This resulted in 24 different waste-type groups encompassing 133 of the 149 tanks, and 16 single tanks. Reference [2] was used to determine the primary and secondary waste types for each tank.

The HTCE model is based on similar principles, but provides tank inventory estimates for all 149 single shell tanks. Here, a chemical composition is specified for each waste type. Then tanks are identified by their volume percentages of all possible waste types, taking into account their process histories. The chemical compositions are then averaged together, weighted by volume, to get a final waste composition estimate for each tank. The HTCE Model does not come up with explicit groups of tanks, as does the SORWT Model.

For this report, a statistical analysis technique called Hierarchical Cluster Analysis [5] was used to identify groups of similar tanks, based on both process history models and sampling data estimates. In this clustering technique, each tank represents a data point. The location of the tank data point in multi-dimensional space is identified by a vector of constituent concentrations. The Euclidean distance between tank data points is calculated. The cluster analysis algorithm finds tanks that are close to each other in the multi-dimensional space and puts them together in clusters or groups. The algorithm starts with each tank in a separate cluster and then groups the two closest clusters into one cluster. This continues until all tanks are within a single group. It is the role of the analyst to determine the number of groups that are appropriate. The Calinski-Harabaz (CH) statistic [5] and MANOVA [9] can be used to aid in determining the optimal number of clusters.

One useful result of the cluster analyses presented in this report is to provide an evaluation of the HTCE and SORWT Models. To this end, all cluster analysis results are provided with the appropriate tank name and the associated SORWT designation and HTCE waste designation. In this section, a cluster analysis using the HTCE estimates, as well as a cluster analysis on the Hanford Defined Wastes, is presented.

2.1 HTCE and SORWT Comparisons

A cluster analysis was performed using, as data, the estimates that are provided in the HTCE study. Because this analysis is based on historical estimates rather than on analytical data, one would expect that the clusters or groups would be fairly well defined, based on the designated primary and secondary waste types as identified by the HTCE study.

The HTCE estimates are identical for several tanks. This is because only one waste type was identified for those tanks, and the chemical composition for that waste type was therefore assigned to each tank. These groups are identified in Table 1. The groups will always remain together in the cluster analyses that follow since the model estimates are identical.

The analytes used in this cluster analysis and in all others that follow are listed in Table 2. WHC identified these analytes as key constituents in the effort to group the Hanford waste tanks.

For this analysis, no transformations were made to the data, although several were considered but determined to be of no significant value. Zero values were not altered. It was decided to create 40 clusters,

Table 1: Tanks With Identical Estimates from HTCE Model

HTCE Waste Type	Tank
UR/TBP	TY-104 TY-105 TY-106
1C44-51/CW	BX-107 BX-108 C-110 T-107 TX-109
MW44-51	BX-104 BX-105 BX-106 C-201 C-202 C-203 C-204 U-101 U-104
CWR/A152-60	U-201 U-202 U-203 U-204
224	B-201 B-202 B-203 B-204 T-201 T-202 T-203 T-204
P'63-67	A-105 AX-104

Table 2: Selected Analytes for Tank Grouping Study

Aluminum	Bismuth
Iron	Sodium
Cesium-137	Strontium-90
Uranium	Plutonium
Phosphate	Manganese
Chromium	Nitrite
Nitrate	Fluoride
Percent Water	Silicon

based solely on the fact that there are 40 clusters or tank groups identified in the SORWT Model, and to facilitate comparisons between models. The results of the analysis are given in Table 3, along with the SORWT designations and the HTCE designations for comparison. The List of Terms provides definitions for these waste type designations.

Table 3: Cluster Analysis Results for HTCE Data

Cluster	Tank	SORWT			HTCE	
		Group	Primary	Secondary	Primary	Secondary
1	AX-102	23	CCPLX	DSSF	A1-SC	B
1	AX-103	23	CCPLX	DSSF	A1-SC	P'63-67
2	S-107	1	R	EB	R'52-58	T2-SC
2	SX-101	1	R	EB	R'52-58	S1-SC
3	SX-109	4	R		R-SC	R'52-58
3	SX-114	4	R		R-SC	R'59-67
4	TX-102	1	R	EB	T2-SC	MW44-51
4	TX-103	21	TBP	EB	T2-SC	MW44-51
4	TX-104	1	R	EB	T2-SC	MW44-51
4	TX-105	1	R	EB	T2-SC	MW52-56
4	TX-106	1	R	EB	T2-SC	R'52-58
4	TX-107	1	R	EB	T2-SC	MW52-56
4	TX-108	35	EB	DW	T2-SC	UR/TBP
4	TX-110	2	EB	1C	T2-SC	1C52-56/CW
4	TX-111	2	EB	1C	T2-SC	1C52-56/CW
4	TX-112	2	EB	1C	T2-SC	T1-SC
4	TX-114	2	EB	1C	T2-SC	T1-SC
4	TX-115	9	EB	R	T2-SC	UR/TBP
4	TX-118	6	EB	CW	T2-SC	Z/PFP
5	BX-109	5	TBP	CW	UR/TBP	1C44-51/CW
5	C-108	13	TBP-F	1C	UR/TBP	1C44-51/CW
5	TY-104	22	TBP	1C-F	UR/TBP	
5	TY-105	37	TBP		UR/TBP	
5	TY-106	38	TBP	DIA	UR/TBP	
6	C-102	30	CW	TBP	CWP/Al61-72	CWP/Al56-60
6	C-105	32	TBP	SR-WASH	CWP/Al56-60	CWP/Al61-72
6	T-101	17	CW	MIX	CWR/Al61-67	MW44-51
6	T-102	17	CW	MIX	CWP/Al61-72	MW44-51
6	T-103	17	CW	MIX	CWP/Al61-72	T2-SC
7	B-104	29	2C	EB	1C44-51/CW	2C44-51
7	B-107	8	1C	EB	1C44-51/CW	
7	BX-107	12	1C	TBP	1C44-51/CW	
7	BX-108	5	TBP	CW	1C44-51/CW	
7	BX-110	8	1C	EB	1C44-51/CW	BY-SC
7	BX-112	8	1C	EB	1C44-51/CW	B-SC
7	C-107	10	1C	CW	1C44-51/CW	SRR/FMJ
7	C-110	12	1C	TBP	1C44-51/CW	
7	T-104	33	1C		1C44-51/CW	1C52-56/CW
7	T-107	10	1C	CW	1C44-51/CW	
7	TX-109	2	EB	1C	1C44-51/CW	
8	C-109	13	TBP-F	1C	TFeCN	1C44-51/CW
8	C-112	13	TBP-F	1C	TFeCN	CWP/Al61-72
9	B-112	16	2C	5-6	BY-SC	2C44-51

Table 3: Cluster Analysis Results for HTCE Data

Cluster	Tank	SORWT			HTCE	
		Group	Primary	Secondary	Primary	Secondary
9	TY-103	22	TBP	1C-F	UR/TBP	T2-SC
10	C-111	13	TBP-F	1C	1C44-51/CW	CWP/A156-60
10	U-110	10	1C	CW	1C52-56/CW	MW44-51
10	U-112	40	UK		1C44-51/CW	R'52-58
11	S-101	1	R	EB	R'52-58	S1-SC
11	S-110	1	R	EB	S1-SC	R'52-58
11	SX-104	1	R	EB	S1-SC	R'52-58
12	B-102	6	EB	CW	B-SC	MW44-51
12	B-103	6	EB	CW	B-SC	MW44-51
12	B-105	2	EB	1C	B-SC	1C44-51/CW
12	B-106	12	1C	TBP	B-SC	
12	B-108	8	1C	EB	B-SC	1C44-51/CW
12	B-109	8	1C	EB	B-SC	CWP/A161-72
12	T-108	12	1C	TBP	T1-SC	1C44-51/CW
12	T-109	21	TBP	EB	T1-SC	
13	S-102	1	R	EB	S2-SS	S1-SC
13	S-103	1	R	EB	S1-SC	S2-SS
13	SX-106	1	R	EB	S2-SS	S1-SC
13	U-103	9	EB	R	S1-SC	S2-SS
13	U-105	6	EB	CW	S2-SS	T2-SC
13	U-107	6	EB	CW	S2-SS	MW44-51
13	U-108	6	EB	CW	S1-SC	S2-SS
13	U-109	6	EB	CW	S1-SC	S2-SS
13	U-111	9	EB	R	S2-SS	S1-SC
14	TX-101	34	R	MIX	R'52-58	T2-SC
14	S-104	4	R		R'52-58	CWR/A161-67
14	SX-113	24	R	DIA	R'52-58	
15	BX-111	8	1C	EB	BY-SC	1C44-51/CW
15	BY-101	3	TBP-F	EB-ITS	BY-SC	MW44-51
15	BY-102	19	TBP	EB-ITS	BY-SC	MW44-51
15	BY-103	3	TBP-F	EB-ITS	BY-SC	CWP/A161-72
15	BY-106	3	TBP-F	EB-ITS	BY-SC	PFeCN/1
15	BY-109	19	TBP	EB-ITS	BY-SC	MW44-51
15	BY-111	3	TBP-F	EB-ITS	BY-SC	MW44-51
15	BY-112	3	TBP-F	EB-ITS	BY-SC	MW44-51
16	T-111	15	2C	224	2C52-56	2C44-51
16	T-112	15	2C	224	2C52-56	
17	SX-111	4	R		R'52-58	R'59-67
17	SX-112	4	R		R'52-58	R'59-67
18	TX-113	2	EB	1C	T2-SC	1C44-51/CW
18	TX-116	2	EB	1C	T1-SC	T2-SC
18	TX-117	2	EB	1C	T2-SC	T1-SC
18	TY-102	2	EB	1C	T2-SC	T1-SC
19	B-110	16	2C	5-6	2C44-51	P'63-67
19	B-111	16	2C	5-6	2C44-51	P'63-67
19	T-105	10	1C	CW	2C44-51	1C52-56/CW
19	T-110	15	2C	224	2C44-51	2C52-56
20	S-105	1	R	EB	S1-SC	R'52-58

Table 3: Cluster Analysis Results for HTCE Data

Cluster	Tank	SORWT			HTCE	
		Group	Primary	Secondary	Primary	Secondary
20	S-106	1	R	EB	S1-SC	CWR/A161-67
20	S-108	1	R	EB	S1-SC	R'52-58
20	S-109	1	R	EB	S1-SC	CWR/A161-67
20	S-111	1	R	EB	S1-SC	R'52-58
20	S-112	1	R	EB	S1-SC	R'52-58
20	SX-102	1	R	EB	S1-SC	S2-SS
20	SX-103	1	R	EB	S1-SC	R'52-58
20	SX-105	1	R	EB	S1-SC	R'52-58
20	U-106	9	EB	R	S1-SC	MW44-51
21	BY-104	3	TBP-F	EB-ITS	BY-SC	PFeCN/2
21	BY-105	3	TBP-F	EB-ITS	BY-SC	PFeCN/2
21	BY-107	3	TBP-F	EB-ITS	BY-SC	PFeCN/1
21	BY-110	3	TBP-F	EB-ITS	BY-SC	PFeCN/1
22	A-101	11	DSSF	NCPLX	A1-SC	A2-SS
22	A-102	11	DSSF	NCPLX	A1-SC	A2-SS
22	A-103	11	DSSF	NCPLX	A1-SC	AR
22	AX-101	11	DSSF	NCPLX	A1-SC	A2-SS
23	BX-102	5	TBP	CW	MW44-51	CWP/A161-72
23	BX-103	5	TBP	CW	MW44-51	CWP/A161-72
23	BX-104	5	TBP	CW	MW44-51	
23	BX-105	5	TBP	CW	MW44-51	
23	BX-106	5	TBP	CW	MW44-51	
23	C-201	14	HS		MW44-51	
23	C-202	14	HS		MW44-51	
23	C-203	14	HS		MW44-51	
23	C-204	14	HS		MW44-15	
23	U-101	4	R		MW44-51	
23	U-104	24	R	DIA	MW44-51	
24	SX-107	4	R		R'52-58	R'59-67
24	SX-108	4	R		R'52-58	R'59-67
25	SX-110	4	R		R'59-67	R-SC
25	SX-115	4	R		R'59-67	R-SC
26	U-201	18	CW		CWR/A152-60	
26	U-202	18	CW		CWR/A152-60	
26	U-203	18	CW		CWR/A152-60	
26	U-204	39	R	2C	CWR/A152-60	
27	B-201	7	224		224	
27	B-202	7	224		224	
27	B-203	7	224		224	
27	B-204	7	224		224	
27	T-201	7	224		224	
27	T-202	7	224		224	
27	T-203	7	224		224	
27	T-204	7	224		224	
28	A-105	26	P	IX	P'63-67	
28	AX-104	28	EVAP	NCPLX	P'63-67	
29	A-104	25	SLUICE	P	AR	P'56-62
30	A-106	27	CCPLX	NCPLX	A1-SC	A2-SS

Table 3: Cluster Analysis Results for HTCE Data

Cluster	Tank	SORWT			HTCE	
		Group	Primary	Secondary	Primary	Secondary
31	B-101	6	EB	CW	B-SC	BL
32	BX-101	5	TBP	CW	MW44-51	BL
33	BY-108	3	TBP-F	EB-ITS	PFeCN/1	BY-SC
34	C-101	5	TBP	CW	UR/TBP	CWP/A161-72
35	C-103	20	SRS	SR-WASH	CWP/A161-72	AR
36	C-104	31	CW	OWW	CWP/A161-72	SRR/FMJ
37	C-106	20	SRS	SR-WASH	AR	BL
38	T-106	10	1C	CW	1C52-56/CW	CWR/A152-60
39	TY-101	36	1C-F	EB	1CFeCN	T1-SC
40	U-102	9	EB	R	T1-SC	S2-SS

Table 3 is summarized in Figure 1. This figure gives the number of tank counts that are contained in each HTCE cluster by SORWT group combination. The tank count totals shown in the margins are given for both SORWT and HTCE models. The dashed lines in the plot divide the clusters or groups containing more than one tank from those with single tanks. A square around a tank count in the plot indicates that the HTCE cluster is exactly the same as a SORWT group (i.e., a perfect match). An octagon around a tank count in the plot indicates an entire SORWT group is contained in a larger HTCE cluster. A diamond around a tank count indicates that an entire HTCE cluster is contained in a larger SORWT group.

Several general observations can be inferred from Table 3 and Figure 1 as follows:

1. The general pattern seems to be that the clusters are determined by the primary waste type as indicated by HTCE. The primary waste type is defined, for purposes of this paper, to be the waste type that comprises the largest volume amount in the tank. While the general rule is for clusters to be based on primary waste type, there are usually several tanks for each primary waste type that do not follow this pattern.
2. This analysis provides an interesting comparison to the SORWT waste designations, which generally agree with the HTCE designations. However, it must be noted that some conventions for identifying the wastes were different for HTCE than for SORWT. For example, SORWT has one "generic" designation for 1C wastes, while HTCE has two separate designations for 1C wastes.
3. The results of the cluster analysis show that the HTCE and SORWT Models grouped tanks identically for 18 of the 149 single-shell tanks. This can be seen by summing the number of tank counts that are contained in squares in Figure 1.
4. The cluster analysis also shows that the HTCE and SORWT models group tanks somewhat similarly for 55 of the 149 tanks. This was calculated by summing the tank counts from Figure 1 that are contained in any symbol with the exception of the single tanks contained in octagon and diamond shaped symbols. This result indicates that there is a measure of disagreement between the HTCE and SORWT models (i.e., disagreement between model groupings for 63% of the tanks). As can be observed in Figure 1, only 37% of the tanks correspond directly between the two grouping levels. However, other intersections between the two grouping methods become apparent (i.e., SORWT Group 1 and HTCE Clusters 4 and 20). This information can be used to guide future sampling efforts and data interpretation.

Some specific observations regarding the cluster analysis are as follows.

1. For tank BX-108, SORWT does not list 1C waste at all, while HTCE lists 1C waste as the only waste type.

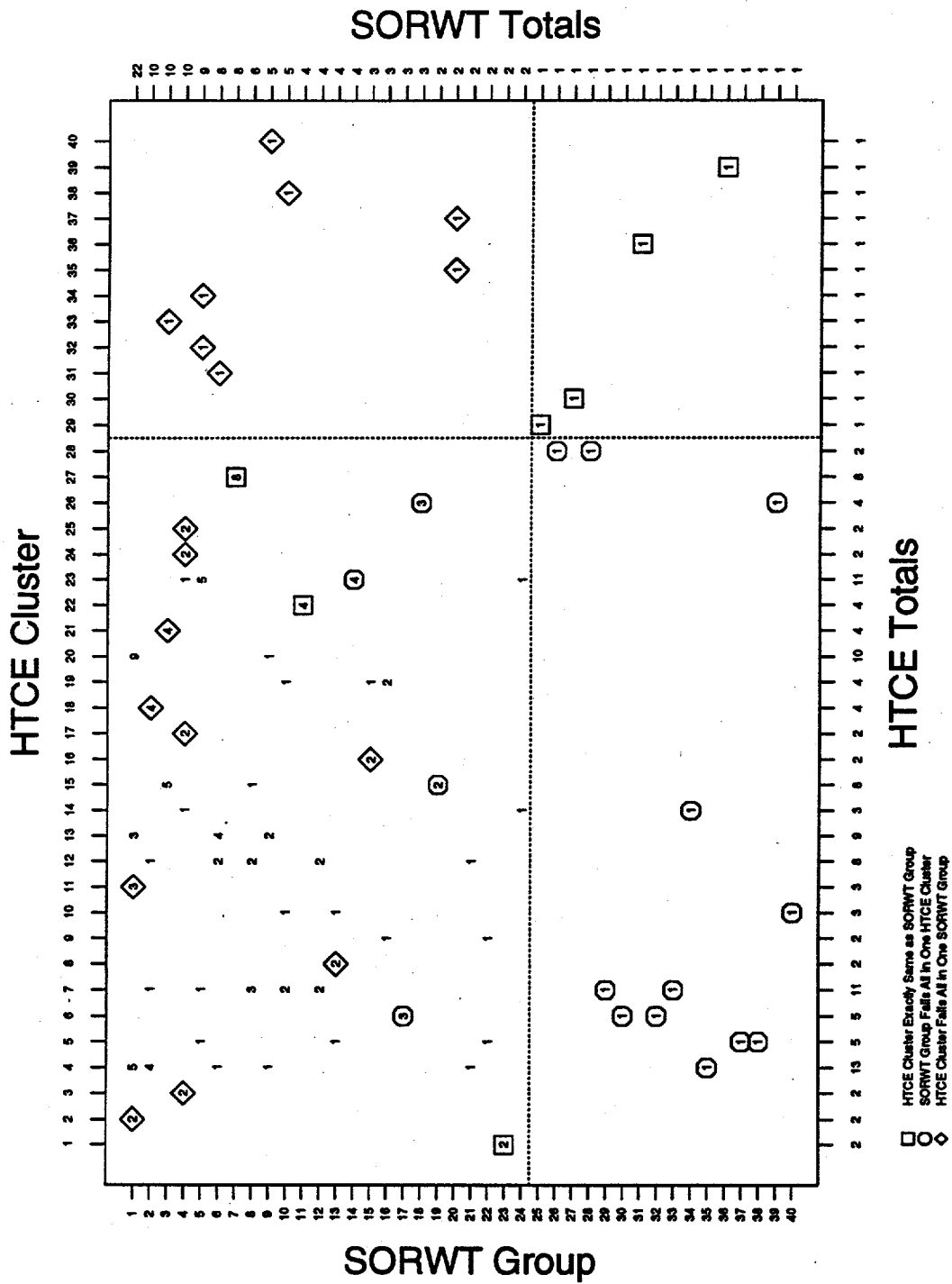


Figure 1: Summary of HTCE Cluster Analysis

2. The SORWT groups comprised of 2 or more tanks that were preserved were groups 17, 19, 11, 14, 18, 23, and 7, thus showing good agreement with the HTCE Model. These SORWT groups contain 3, 2, 4, 4, 3, 2 and 8 tanks respectively.
3. Tank B-107 has one waste type (1C44-51/CW) listed by HTCE, but the data that was obtained for the analysis was not identical with the other tanks that have the same single waste type. This is inconsistent with the way that the HTCE estimates were supposed to be obtained.
4. Salt cake (SC) designations in HTCE are the same as evaporator bottom (EB) designations in SORWT. Note that for Cluster 20, SORWT and HTCE identify the same two largest waste types based on volume; however, the HTCE identifies a salt cake waste (S1-SC) as the primary waste type and SORWT identifies it (EB Waste) as the secondary waste type (Tank U-106 is an exception in this cluster, since both models agree on the primary waste type).

2.2 Hanford Defined Wastes Cluster Analysis

A cluster analysis was also conducted on the 50 Hanford Defined Waste types used to obtain the HTCE estimates. This cluster analysis identifies groups of waste types that have similar waste characteristics. Table 4 shows the results of the cluster analysis. Concentrations of the analytes in Table 2 were used once again to determine groups of waste types.

Table 4: Hanford Defined Wastes Cluster Analysis Results

Group	Hanford Defined Wastes
1	T2-SC BY-SC
2	UR-TBP PL
3	CWR50-61 CWP/A156-60
4	CSR DC
5	SRR CC1
6	TH66 TH70 PL83-88
7	S1-SC A1-SC
8	S2-SS A2-SS
9	PFeCN1 PFeCN2
10	CW/Zr66-72 CWP/Zr83-88
11	DW N
12	OWW55-65 OWW69-72
13	B-SC T1-SC

The CH statistic shows that there are 36 groups in the Hanford Defined Waste estimates. Twenty-three of these groups are single waste types. The other 13 groups shown in Table 4 identify waste types that are very similar, based on the analytes used in this cluster analysis. The waste types within a given group could potentially be merged into a new single waste type.

This cluster analysis also revealed two more groups of similar waste types that are discussed later in this report. These similar groups are the following:

1. 2C44-51, 1C44-51/CW, and 1C52-56/CW wastes
2. 2C52-56 and 224 wastes.

The similarities between these waste types are not as strong as those in Table 4 but they are worth noting. These five waste types were left as individual tanks in this cluster analysis. If the cluster tree were cut to form 25 clusters rather than 36, these five waste types would form the two groups listed above.

3 Tank Grouping Efforts Based on Analytical Data

This section presents the results from three tank grouping analyses based on analytical data. A brief description of the analytical data gathering effort is given prior to the discussion of the tank grouping analyses. The first tank grouping analysis is based on TCR mean concentration estimates. The second analysis is based on TCR spatial variability estimates. The final tank grouping analysis is based on mean concentration estimates from all sampling data gathered (from TCR, TWINS, and Braun data sources).

3.1 Analytical Data Description

This section gives a brief description of the analytical data that was used in the tank grouping studies. A more detailed discussion of the data sources, the editing of analytical results from those sources, and data handling information is contained in Appendix A.

Tank waste samples are typically analyzed to estimate concentration or properties of a large suite of chemical, physical and radiological constituents. For the purposes of this tank grouping study, a small subset of this suite of constituents was selected, as listed in Table 2. These analytes were selected because they are expected to be key in the effort to distinguish between different groups of similar tanks.

Several sources of analytical data were considered for input into a tank grouping study, including:

1. Tank Characterization Reports (TCR) - The estimates taken from the TCRs are generally based on relatively recent (1990 to present) core sampling activities. These estimates are considered to be the most reliable sample based estimates of tank contents.
2. Tank Waste Information Network System (TWINS) - This network contains a large number of analytical results from recent tank sampling activities.
3. Kaiser Dataset - This dataset contains a large number of analytical results from many sampling activities. The results from this dataset had not been released by ICF Kaiser during the time of the data gathering effort.
4. Braun Dataset - This dataset also contains a large number of analytical results from many tank sampling activities. The quality of this data source is somewhat less than the other data sources.

These different data sources vary in data quality and are not necessarily mutually exclusive.

The goal of the data gathering effort was to assemble a set of data as typical as possible of the waste that exists in the single-shell tank population at Hanford.

The TCR estimates³, TWINS data and Braun data sources were selected for the following reasons:

1. These three data sources collectively represent a majority of the analytical results from Hanford condensed phase samples.
2. All of these sources were available in early FY95 when the data gathering effort was carried out.
3. Analytical results across several tanks were relatively easy to retrieve from the data sources.

Figure 2 illustrates how the tank grouping dataset was assembled using these three sources. The tank characterization reports were used as the first source, since they are expected to be the most "typical" set of analytical estimates available. However, TCR estimates were available for only 15 tanks at the time the data was gathered. If there were no TCR estimates available for a given tank, the TWINS database was used as the second source in the data search. Finally, if TCR and TWINS analytical information was not available for a given tank, the Braun data was used. The Braun dataset is quite extensive and contains analytical results taken throughout the service life of a tank and after its removal from service, for more than 100 of the 177 tanks at Hanford. However, this dataset has several issues with regard to the data quality. There are some incomplete sample records and incorrect sample results reported in the Braun dataset. Appendix

³The TCR estimates are mean concentration estimates calculated from analytical results, whereas the TWINS and Braun data are raw analytical results

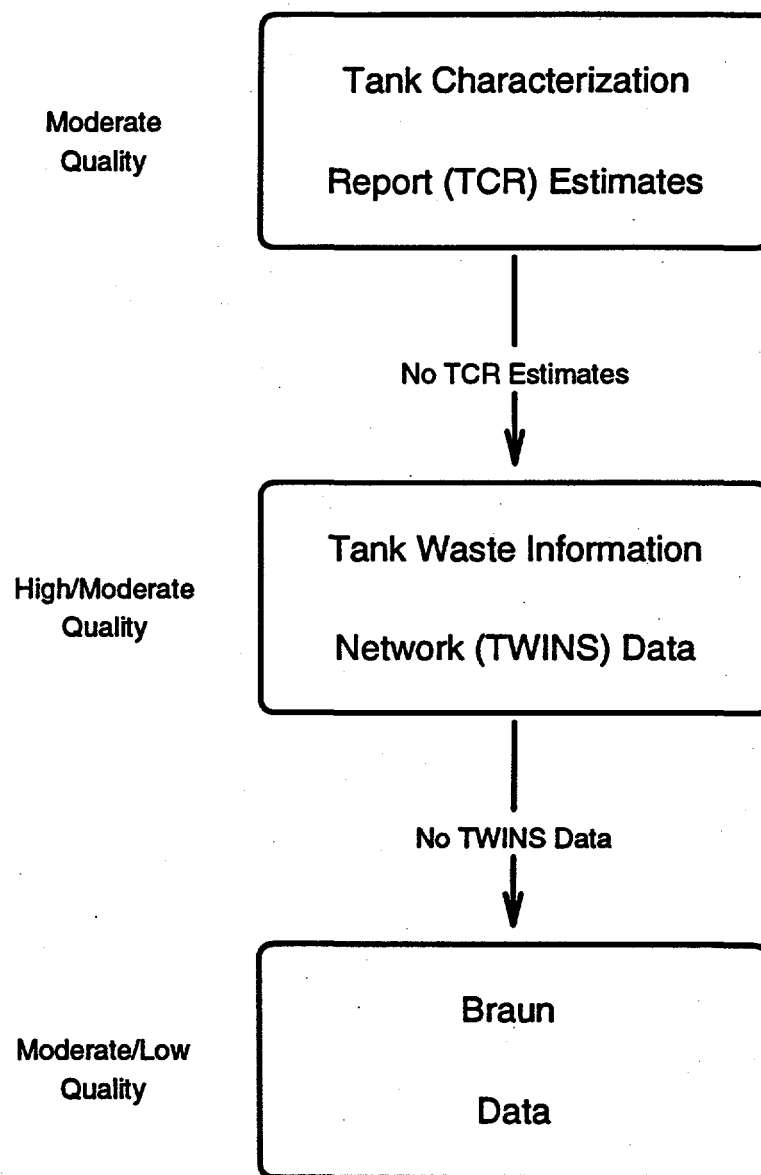


Figure 2: Tank Grouping Data Collection

A discusses how these problems with the Braun dataset were dealt with. Because of these problems, the Braun dataset was used as the last data source for the tank grouping studies.

Once the tank grouping dataset was put together using these three data sources, a simple average was calculated of all the analytical results for each tank and analyte combination. These vectors of means are the data inputs into the tank grouping study discussed in this report.

No analytical data was available for some tank and analyte combinations. Rather than eliminate the tanks with missing values from the analysis, a procedure called the EM algorithm [4] was used to estimate the missing values. These estimated missing values are used only to establish tank groups and are not used in any tank inventory or spatial variability estimate calculations for the tank groups identified.

Appendix A contains a more complete discussion of the steps that were taken to obtain a dataset appropriate for the tank grouping studies. Also included in Appendix A is a detailed discussion of the EM algorithm used to estimate the missing mean values.

Appendix A also contains tank mean plots for each constituent used in the tank grouping study.

3.2 Tank Grouping Analysis Based on TCR Mean Concentration Estimates

The mean concentration estimates from the TCRs are generally accepted as the "best" or most reliable tank characterization estimates based on analytical data. It was noted in Section 3.1 that the quality of sample results from the Braun dataset was moderate to low. For these two reasons, a tank grouping analysis was conducted with the TCR mean concentration estimates alone. Only 15 tanks were used in this analysis. Despite the small number of tanks, the potential for finding meaningful patterns in the data is still good, since the TCR estimates are more reliable (smaller absolute bias and uncertainty than other data).

The analytical dataset for these 15 tanks was nearly complete. Only two values had to be estimated using the EM algorithm.

3.2.1 TCR Cluster Analysis Using Normalized Means

The first cluster analysis was run so that all of the constituents listed in Table 2 had equal weight in determining group patterns (i.e., the means were normalized). In this analysis, constituents found at trace levels (e.g., chromium and manganese) had the same influence in determining tank groups as did the large weight % constituent contributors (e.g., water and sodium). The mean concentrations for each constituent (across all tanks) were normalized to make the equal weighting of constituents possible. The results of this cluster analysis are useful if all constituents found in Table 2 are equally important.

Table 5 contains the results of this cluster analysis, showing the clusters or groups of "like" tanks based on the TCR constituent mean concentrations. The SORWT and HTCE primary and secondary wastes types are given. Note that for this set of 15 tanks the two models are in general agreement. The Calinski-Harabaz (CH) statistic [5] and MANOVA [9] were used to aid in the effort to determine how many "significantly" different clusters exist in the data.

Based on the two statistics just mentioned, 10 clusters were found in the data (at the 95% confidence level). Tank Group 1 is made up of five tanks which received primarily first cycle (1C) and second cycle (2C) waste from the bismuth phosphate process. The only other group that is made up of more than one tank is Tank Group 5. Based on the process history models, 224 Waste from the bismuth phosphate process is the waste type common to these two tanks.

Another meaningful tank grouping that did not provide as much separation between the tank groups as did the 10 clusters, is a cut at 5 clusters in the data. The MANOVA indicated that there were significant differences between the 5 clusters at the 95% confidence level. Note that two more 1C/2C waste tanks are added to Tank Group 1 from the 10-cluster analysis to form the new Tank Group 1 for the 5-cluster analysis. This relatively strong grouping of 1C and 2C waste type tanks suggests that there may not be a significant difference between these two waste types. The HTCE Model separates 1C and 2C wastes further into two waste types for each, based on a bismuth phosphate process change in 1952. The 2C44-51, 1C44-51, and the 1C52-56 waste types show similar patterns in the cluster analysis. This close relationship between these waste types was previously noted by staff from LANL and WHC who were involved in the development of the process history models. Note that Tank T-111, which primarily contains 2C52-56 and smaller

Table 5: Cluster Analysis Results for Normalized TCR Mean Concentration Estimates

Clusters		Tank	SORWT			HTCE			Cluster Distinguishing Characteristics
5	10		Grp#	Primary	Secondary	Grp#	Primary	Secondary	
1	1	B-110	16	2C	5-6	19	2C44-51	P'63-67	High NO_3
1	1	B-111	16	2C	5-6	19	2C44-51	P'63-67	Low Mn
1	1	BX-107	12	1C	TBP	7	1C44-51/CW		High Na
1	1	C-110	12	1C	TBP	7	1C44-51/CW		High F
1	1	T-104	33	1C		7	1C44-51/CW	1C52-56/CW	High PO_4
1	2	T-107	10	1C	CW	7	1C44-51/CW		Low Bi
1	3	U-110	10	1C	CW	10	1C52-56/CW	MW44-51	
2	4	B-201	7	224		27	224		Low Sr^{90} , NO_2
2	5	B-202	7	224		27	224		Low U , Cs^{137}
2	5	T-111	15	2C	224	16	2C52-56	2C44-51	High Cr , Mn
3	6	C-109	13	TBP-F	1C	8	TFeCN	1C44-51/CW	High Sr^{90} , NO_2
3	7	C-112	13	TBP-F	1C	8	TFeCN	CWP/Al61-72	High U , Cs^{137}
									Low Cr , Mn
4	8	T-102	17	CW	MIX	6	CWP/Al61-72	MW44-51	Unclear
4	9	T-105	10	1C	CW	19	2C44-51	1C52-56/CW	
5	10	S-104	4	R		14	R'52-58	CWR/Al61-72	Low PO_4
									High Al
									Low F , H_2O , Fe

quantities of 224 wastes, does not group with the other 1C and 2C waste tanks. Tank Groups 2, 3 and 5 (from a cut at 5 clusters) also show good agreement with the process history models. Tank Group 4 is the only group that does not show strong agreement with either of the process history models at the 5 cluster cut.

Note that Tanks B-201, B-202 and T-111 are grouped together with a cut at five clusters. With a cut at ten clusters, Tank B-201 becomes its own group. This may be due to the abnormally large bismuth concentration for Tank B-201. The Hanford Defined Waste cluster analysis discussed in Section 2 indicated that there are similarities between 2C52-56 waste and 224 waste. Tanks B-201, B-202 and T-111, which contain these waste types, are also grouping together based on the TCR analytical data. These results suggest that these waste types are very similar, and may be potential candidates for merging into a single waste type. It would be useful to have analytical results for other tanks with 2C52-56 waste as the primary waste type, to confirm these results.

The last column in Table 5 identifies the distinguishing characteristics of the 5-cluster analysis. A principal components analysis (PCA) (see Reference [9]) was conducted to identify these characteristics. For example, Tank Group 1 separated from the other clusters of tanks due to the generally high concentrations of nitrate, sodium, fluoride, and phosphate that are found in 1C/2C wastes. Also, low concentrations of manganese and bismuth contributed to the separation of these tanks from the other tanks in this cluster analysis.

As mentioned earlier, the 10 tank groups provided a stronger separation of tanks into "like" groups. The strongest result of this cluster analysis is therefore that the TCR estimates generally indicate that 1C and 2C waste types may not be significantly different, and tanks containing these two waste types may be grouped together.

3.2.2 TCR Cluster Analysis Using Raw Means

A second cluster analysis was conducted using the same TCR mean concentration estimates. But this cluster analysis weighted the contribution of each constituent to the determination of tank groups proportionally to

Table 6: Cluster Analysis Results for TCR Mean Concentration Estimates (Constituents Weighted According to Magnitude)

Cluster	Tank	SORWT			HTCE			Cluster Distinguishing Characteristics
		Grp#	Primary	Secondary	Grp#	Primary	Secondary	
1	C-109	13	TBP-F	1C	8	TFeCN	1C44-51/CW	Low H_2O High Al High Na
1	S-104	4	R		14	R'52-58	CWR/Al61-72	
1	T-102	17	CW	MIX	6	CWP/Al61-72	MW44-51	
1	U-110	10	1C	CW	10	1C52-56/CW	MW44-51	
2	B-110	16	2C	5-6	19	2C44-51	P'63-67	Medium/High H_2O Medium/Low Al Medium/High NO_3 Medium/High Na
2	B-111	16	2C	5-6	19	2C44-51	P'63-67	
2	B-201	7	224		27	224		
2	BX-107	12	1C	TBP	7	1C44-51/CW		
2	C-110	12	1C	TBP	7	1C44-51/CW		
2	C-112	13	TBP-F	1C	8	TFeCN	CWP/Al61-72	
2	T-107	10	1C	CW	7	1C44-51/CW		
3	B-202	7	224		27	224		High H_2O Low Al (except T-105) Low NO_3 Low Na
3	T-104	33	1C		7	1C44-51/CW	1C52-56/CW	
3	T-105	10	1C	CW	19	2C44-51	1C52-56/CW	
3	T-111	15	2C	224	16	2C52-56	2C44-51	

its concentration. In this analysis, constituents found in large concentration, such as sodium, have greater weight than constituents such as manganese, chromium, and radionuclides that are found at trace levels.

The results of this cluster analysis, Table 6, show three significantly different groups of similar tanks at the 95% confidence level, based on the CH statistic and MANOVA. A different grouping of tanks is obtained using this second approach. However, one pattern is consistent with the previously reported cluster analysis. Note that Tank Group 2 in Table 6 is somewhat similar to Tank Group 1 in Table 5. Tank Group 2 is primarily composed of 1C and 2C waste tanks. Generally these 1C/2C waste tanks group together regardless of how the constituents are weighted (i.e., normalized or weighted proportional to concentration). Note that the constituents that drive the cluster analysis are the larger weight percent contributors. Tank Group 2 is separated from the other two groups in Table 6 due to medium to high levels of water, sodium, and nitrate; and medium to low levels of aluminum, as compared to the other tanks considered in this study.

Note that Tank T-102 appears in Tank Group 1. Low concentrations of water is the largest factor that separates this group from the other two. The T-102 TCR noted that the core samples taken near the edges of the tank were dryer than the rest of the tank waste. It is possible that Tank T-102 may have grouped with Tank Groups 2 or 3 if the water results for this tank were more representative (higher).

The results of this cluster analysis add further strength to the conclusion that 1C/2C wastes might be considered as a single waste type group.

3.3 Tank Grouping Analysis Based on TCR Spatial Variability Estimates

Another cluster analysis was based on the TCR spatial variability estimates expressed as relative standard deviations (RSD). These estimates were obtained from the TCRs. In several instances, spatial variability estimates were not available, for reasons which include incomplete core recovery or the report of less than the detection limit for one of the cores. Analytical results from at least two core samples are needed to estimate spatial variability. For several of the constituents, there were multiple estimates of spatial variability, i.e., one for each sample preparation technique used. The fusion sample preparation method was used for the elementals metal and radionuclides.

The spatial variability estimates were derived from core-to-core differences. The constituents listed in

Table 7: Cluster Analysis Results for the TCR Relative Standard Deviations

Five Clusters	Ten Clusters	Tank	SORWT			HTCE		
			Grp#	Primary	Secondary	Grp#	Primary	Secondary
1	1	B-110	16	2C	5-6	19	2C44-51	P'63-67
1	1	B-111	16	2C	5-6	19	2C44-51	P'63-67
1	1	BX-107	12	1C	TBP	7	1C44-51/CW	
1	1	T-104	33	1C		7	1C44-51/CW	1C52-56/CW
1	2	B-202	7	224		27	224	
1	3	T-111	15	2C	224	16	2C52-56	2C44-51
2	4	C-110	12	1C	TBP	7	1C44-51/CW	
2	5	S-104	4	R		14	R'52-58	CWR/Al61-72
3	6	C-109	13	TBP-F	1C	8	TFeCN	1C44-51/CW
3	7	C-112	13	TBP-F	1C	8	TFeCN	CWP/Al61-72
3	8	T-105	10	1C	CW	19	2C44-51	1C52-56/CW
3	8	T-107	10	1C	CW	7	1C44-51/CW	
4	9	B-201	7	224		27	224	
5	10	U-110	10	1C	CW	10	1C52-56/CW	MW44-51

Table 2 were used for this analysis, with the exception of strontium-90 and water. These were dropped because spatial variability estimates of these constituents were not available for several of the tanks. Tank T-102 was not included in the analysis because one of the cores taken had very poor sample recovery and analytical results were available only for the other one.

Table 7 contains the results of the cluster analysis on TCR spatial variability estimates. The selection of 5 and 10 tank clusters or groups for this analysis is based entirely on the fact that 5 and 10 clusters were specified in the earlier cluster analysis based on TCR mean concentrations. The CH statistic and MANOVA could not be calculated for this analysis, due to the large number of missing values in the data matrix.

One group is notable in Table 7. Cluster 1 from both grouping cuts contains only 1C and 2C waste tanks, with the exception of Tank B-202. This result suggests that the horizontal spatial variability (as estimated from the core sampling data) in tanks containing 1C and 2C waste is similar. This is further support for the merging of these two waste types.

3.4 Tank Grouping Analysis Based on Mean Concentrations of All Analytical Results

As mentioned in Sections 3.1 and 3.2, many of the analytical results from the Braun dataset are not representative of the waste as it currently exists in the Hanford waste tanks, for one or more of the following reasons:

1. The Braun samples may have been taken before a tank was sluiced (emptied).
2. Unit conversions for the sample measurements are occasionally incorrect.
3. Information such as the sampling method used and the location of a sample in the tank may be missing.
4. Sample preparation information may be missing (e.g., water or acid digestion).

This list is not all inclusive. Corrections were made to the sample results if the information needed to make the corrections was available. Other results were deleted from the dataset if they were non-representative of the current conditions in the waste. These exclusions and corrections are described in Appendix A. We

expect that some non-representative sample results still remain in the data, but they cannot be identified due to lack of information about the samples.

This tank grouping effort is expected to continue into the future. The Braun data will be replaced by more reliable core sampling data when it becomes available.

Even in the presence of several potential non-representative sampling results, a cluster analysis that uses Braun sampling results is worthwhile, since the Braun dataset contains sample results for more than 100 of the 149 single-shell tanks. By contrast, there are TCR sampling estimates for only 15 of the single-shell tanks (at the time the data was gathered). Even in the presence of some anomalies, there should be some meaningful grouping patterns in the data.

One objective of these cluster analyses is to see if the 15 single-shell tanks with TCR estimates group similarly (as noted in Section 3.2) in the presence of the "noisier" Braun data.

A global set of data was constructed by supplementing the TCR data with TWINS and Braun data, as explained in Section 3.1. The TCR results, representing 15 tanks, were supplemented with one additional tank from TWINS and 61 tanks from Braun, for a total of 77 tanks. Approximately half of the single shell tanks are represented in this cluster analysis (77 out of 149).

As with the cluster analyses conducted on TCR mean concentrations, two cluster analyses were performed on the global set of analytical data. The first cluster analysis weights the constituents equally. The second cluster analysis weights the constituents proportionally to their mean concentrations.

Principal component analysis is used again to identify which constituents are contributing to the grouping patterns found in the cluster analyses. The group characteristics are listed in the tables in the order of their contribution (from high to low).

3.4.1 Global Cluster Analysis on Normalized Means

Table 8 contains the results for the first cluster analysis. The CH statistic and MANOVA indicate that there are 10 clusters in this global analytical dataset at the 95% confidence level. Despite the lower resolution that exists in this dataset, some meaningful groups can be identified in the table. The fifteen tanks for which there were TCR/TWINS estimates are presented with a bold face font. The following is a summary of these results:

1. Cluster 1 is distinguished from the other groups by its generally high levels of cesium-137, strontium-90, nitrite, and uranium, as well as generally low levels of fluoride and bismuth. Based on the SORWT Model, four of the five tanks in this group contain TBP waste. This result adds credibility to the SORWT Model for this subset of tanks. There is not a clear grouping pattern here based on the HTCE Model. Note that Tanks C-109 and C-112 group together in this analysis as they did in Section 3.2.
2. Tank Cluster 2 is distinguished from the other groups by its generally high levels of phosphate and generally low levels of bismuth and manganese.
3. High levels of bismuth, chromium, and manganese as well as low levels of phosphate distinguish Tank Cluster 5 from the others. Once again, this grouping does not match with either of the process history models.
4. Tank Cluster 6 does not have any clear distinguishing characteristics. This cluster does contain several of the tanks with TCR means (indicated by the bold face font). Tanks B-110, BX-107, C-110, and T-104 containing primarily 1C/2C waste continue to group together (as in Section 3.2) in the presence of tanks with Braun data. However Tank B-111, which grouped with these tanks in Section 3.2, does not group with these tanks in the presence of tanks with Braun data. This indicates that there is a group of tanks with Braun data that group more tightly around the B-111 values than do Tanks B-110, BX-107, C-110 and T-104. It is also interesting to note that Tanks B-202, T-111 and other tanks containing 224 waste are included in Tank Cluster 6.
5. Note that there are several Braun data tanks in Clusters 4 and 6 that contain saltcakes and sludges that are drastically different than the tanks with TCR estimates. This may be attributed to the

"noisy" nature of the Braun data or waste type misclassification by SORWT and HTCE Models or some combination of the two.

6. Tank Cluster 7 is distinguished from the other clusters by its generally high levels of iron, silicon, and strontium-90, as well as generally low levels of nitrate and sodium. Two of the three tanks (i.e., C-103 and C-106) contain strontium wash waste, based on both process history models. This consistency between the process history models and the analytical data is another positive mark for these models.
7. Tank Cluster 9 is the strongest pattern found in this global dataset. It is distinguished from the other tank groups due to its high levels of nitrate, sodium, and nitrite, as well as low levels of iron, silicon, and strontium-90. The high levels of nitrate and sodium signal that these tanks contain large amounts of saltcake. Both of the process history models are in agreement with this Tank Group 9. Based on these models, these tanks primarily contain REDOX wastes and saltcake wastes.
8. Tank Cluster 10 is the single Tank B-201. It is distinguished from the other tanks by its extremely high levels of bismuth (see Figure 2 in Appendix A.5). Additional contributors to the separation of Tank B-201 from the other tanks are its high levels of chromium and manganese and its low levels of strontium-90, nitrite, uranium, cesium-137 and phosphate.
9. Clusters 3, 4, 6 and 8 do not have distinguishing characteristics that are as obvious as the other clusters. This is due to the complicated structure (i.e., high dimensional data) in the data.

Table 8: Cluster Analysis Results for Normalized Global Analytical Dataset

Cluster	Tank	SORWT			HTCE			Cluster Distinguishing Characteristics
		Grp#	Primary	Secondary	Grp#	Primary	Secondary	
1	A-105	26	P	IX	28	P'63-67		High Cs^{137}
1	BX-101	5	TBP	CW	32	MW44-51	BL	High Sr^{90}
1	C-105	32	TBP	SR-WASH	6	CWP/Al56-60	CWP/Al61-72	High NO_2
1	C-109	13	TBP-F	1C	8	TFeCN	1C44-51/CW	High U
1	C-112	13	TBP-F	1C	8	TFeCN	CWP/Al61-72	Low F, Bi
2	B-101	6	EB	CW	31	B-SC	BL	High PO_4
2	BX-104	5	TBP	CW	23	MW44-51		Low Bi
2	BX-112	8	1C	EB	7	1C44-51/CW	B-SC	Low Mn
2	C-104	31	CW	OWW	36	CWP/Al61-72	SRR/FMJ	
2	S-107	1	R	EB	2	R'52-58	T2-SC	
2	T-106	10	1C	CW	38	1C52-56/CW	CWR/Al52-60	
2	T-107	10	1C	CW	7	1C44-51/CW		
2	TY-104	22	TBP	1C-F	4	UR/TBP		
2	TY-105	37	TBP		4	UR/TBP		
2	U-203	18	CW		26	CWR/Al52-60		
3	A-104	25	SLUICE	P	29	AR	P'56-62	Unclear
3	A-106	27	CCPLX	NCPLX	30	A1-SC	A2-SS	
3	SX-115	4	R		25	R'59-67	R-SC	
3	T-105	10	1C	CW	19	2C44-51	1C52-56/CW	
4	A-101	11	DSSF	NCPLX	22	A1-SC	A2-SS	Unclear
4	A-102	11	DSSF	NCPLX	22	A1-SC	A2-SS	
4	A-103	11	DSSF	NCPLX	22	A1-SC	AR	
4	B-111	16	2C	5-6	19	2C44-51	P'63-67	
4	BX-105	5	TBP	CW	23	MW44-51		
4	C-102	30	CW	TBP	25	CWP/Al61-72	CWP/Al56-60	
4	C-204	14	HS		23	MW44-15		
4	S-101	1	R	EB	11	R'52-58	S1-SC	
4	S-102	1	R	EB	13	S2-SS	S1-SC	

Table 8: Cluster Analysis Results for Normalized Global Analytical Dataset

Cluster	Tank	SORWT			HTCE			Cluster Distinguishing Characteristics
		Grp#	Primary	Secondary	Grp#	Primary	Secondary	
4	SX-106	1	R	EB	13	S2-SS	S1-SC	
4	T-204	7	224		27	224		
4	TY-101	36	1C-F	EB	39	1CFeCN	T1-SC	
4	TY-106	38	TBP	DIA	5	UR/TBP		
4	U-105	6	EB	CW	13	S2-SS	T2-SC	
4	U-109	6	EB	CW	13	S1-SC	S2-SS	
5	BY-112	3	TBP-F	EB-ITS	15	BY-SC	MW44-51	Low PO_4 High Bi, Cr, Mn
5	SX-107	4	R		24	R'52-58	R'59-67	
6	B-105	2	EB	1C	12	B-SC	1C44-51/CW	Unclear
6	B-107	8	1C	EB	7	1C44-51/CW		
6	B-110	16	2C	5-6	19	2C44-51	P'63-67	
6	B-202	7	224		27	224		
6	B-203	7	224		27	224		
6	B-204	7	224		27	224		
6	BX-107	12	1C	TBP	7	1C44-51/CW		
6	BX-110	8	1C	EB	7	1C44-51/CW	BY-SC	
6	BY-104	3	TBP-F	EB-ITS	23	BY-SC	PFeCN/2	
6	C-110	12	1C	TBP	7	1C44-51/CW		
6	S-105	1	R	EB	20	S1-SC	R'52-58	
6	S-106	1	R	EB	20	S1-SC	CWR/Al61-72	
6	S-108	1	R	EB	20	S1-SC	R'52-58	
6	S-112	1	R	EB	20	S1-SC	R'52-58	
6	T-104	33	1C		7	1C44-51/CW	1C52-56/CW	
6	T-111	15	2C	224	16	2C52-56	2C44-51	
6	TX-118	6	EB	CW	4	T2-SC	Z/PFP	
6	TY-103	22	TBP	1C-F	9	UR/TBP	T2-SC	
6	U-104	24	R	DIA	23	MW44-51		
6	U-107	6	EB	CW	13	S2-SS	MW44-51	
6	U-108	6	EB	CW	13	S1-SC	S2-SS	
6	U-111	9	EB	R	13	S2-SS	S1-SC	
7	C-103	20	SRS	SR-WASH	35	CWP/Al61-72	AR	Low NO_3 , Na High Fe, Si High Sr^{90}
7	C-106	20	SRS	SR-WASH	37	AR	BL	
7	C-201	14	HS		23	MW44-51		
8	S-104	4	R		14	R'52-58	CWR/Al61-72	Unclear
8	S-110	1	R	EB	11	S1-SC	R'52-58	
8	T-102	17	CW	MIX	6	CWP/Al61-72	MW44-51	
8	U-102	9	EB	R	40	T1-SC	S2-SS	
8	U-110	10	1C	CW	10	1C52-56/CW	MW44-51	
8	U-204	39	R	2C	26	CWR/Al52-60		
9	S-109	1	R	EB	20	S1-SC	CWR/Al61-72	High NO_3 High Na Low Fe Low Si Low Sr^{90} High NO_2
9	S-111	1	R	EB	20	S1-SC	R'52-58	
9	SX-101	1	R	EB	2	R'52-58	S1-SC	
9	SX-102	1	R	EB	20	S1-SC	S2-SS	
9	SX-103	1	R	EB	20	S1-SC	R'52-58	
9	SX-104	1	R	EB	11	S1-SC	R'52-58	
9	SX-105	1	R	EB	20	S1-SC	R'52-58	
9	TY-102	2	EB	1C	18	T2-SC	T1-SC	
9	U-103	9	EB	R	13	S1-SC	S2-SS	
10	B-201	7	224		27	224		Low Cs^{137} , PO_4 Low Sr^{90} , NO_2 , U High Bi, Cr, Mn

3.4.2 Global Cluster Analysis on Raw Means

Table 9 contains the cluster analysis on the global dataset that weights each constituent's contribution proportionally to its concentration. The CH statistic and MANOVA identified three tank groups in the global dataset at the 95% confidence level. The characteristics of each of these groups are as follows:

1. Cluster 1 contains 53 of the 77 tanks in the cluster analysis. This group identifies a large subset of the tanks that are generally high in water and medium to low in nitrate, sodium, and aluminum. This group contains a variety of non-saltcake tanks and several saltcake tanks.
2. Cluster 2 is basically Cluster 9 from the preceding cluster analysis, with the addition of a few more saltcake tanks. The general distinctions between tanks in this group versus tanks in the other groups are the high levels of water, nitrate, and sodium, coupled with low levels of aluminum. This group is almost exclusively made up of saltcake tanks that also contain some REDOX wastes.
3. Tank Cluster 3 is distinguished from the other two groups by generally high levels of aluminum, generally low levels of nitrate and sodium, and medium to low levels of water. Notice that this group has very little saltcake in it, based on the process history models.

Table 9: Cluster Analysis Results for Proportionally Weighted Global Analytical Dataset

Cluster	Tank	Grp#	SORWT		Grp#	HTCE		Cluster Distinguishing Characteristics
			Primary	Secondary		Primary	Secondary	
Group 1 - Sludge and Saltcake Tanks								Low/Medium NO_3 High H_2O Low/Medium Na Low Al
1	A-101	11	DSSF	NCPLX	22	A1-SC	A2-SS	
1	A-102	11	DSSF	NCPLX	22	A1-SC	A2-SS	
1	A-103	11	DSSF	NCPLX	22	A1-SC	AR	
1	A-104	25	SLUICE	P	29	AR	P'56-62	
1	A-105	26	P	IX	28	P'63-67		
1	A-106	27	CCPLX	NCPLX	30	A1-SC	A2-SS	
1	B-105	2	EB	1C	12	B-SC	1C44-51/CW	
1	B-107	8	1C	EB	7	1C44-51/CW		
1	B-110	16	2C	5-6	19	2C44-51	P'63-67	
1	B-111	16	2C	5-6	19	2C44-51	P'63-67	
1	B-201	7	224		27	224		
1	B-202	7	224		27	224		
1	B-203	7	224		27	224		
1	B-204	7	224		27	224		
1	BX-104	5	TBP	CW	23	MW44-51		
1	BX-105	5	TBP	CW	23	MW44-51		
1	BX-107	12	1C	TBP	7	1C44-51/CW		
1	BX-110	8	1C	EB	5	1C44-51/CW	BY-SC	
1	BY-104	3	TBP-F	EB-ITS	21	BY-SC	PFeCN/2	
1	BY-112	3	TBP-F	EB-ITS	15	BY-SC	MW44-51	
1	C-102	30	CW	TBP	6	CWP/Al61-72	CWP/Al56-60	
1	C-103	20	SRS	SR-WASH	35	CWP/Al61-72	AR	
1	C-104	31	CW	OWW	36	CWP/Al61-72	SRR/FMJ	
1	C-106	20	SRS	SR-WASH	37	AR	BL	
1	C-110	12	1C	TBP	7	1C44-51/CW		
1	C-112	13	TBP-F	1C	8	TFeCN	CWP/Al61-72	
1	C-201	14	HS		23	MW44-51		
1	C-204	14	HS		23	MW44-15		
1	S-102	1	R	EB	13	S2-SS	S1-SC	
1	S-105	1	R	EB	20	S1-SC	R'52-58	
1	S-106	1	R	EB	20	S1-SC	CWR/Al61-72	

Table 9: Cluster Analysis Results for Proportionally Weighted Global Analytical Dataset

Cluster	Tank	Grp#	SORWT		Grp#	HTCE		Cluster Distinguishing Characteristics
			Primary	Secondary		Primary	Secondary	
1	S-108	1	R	EB	20	S1-SC	R'52-58	
1	S-112	1	R	EB	20	S1-SC	R'52-58	
1	SX-106	1	R	EB	13	S2-SS	S1-SC	
1	SX-115	4	R		25	R'59-67	R-SC	
1	T-104	33	1C		7	1C44-51/CW	1C52-56/CW	
1	T-105	10	1C	CW	19	2C44-51	1C52-56/CW	
1	T-107	10	1C	CW	7	1C44-51/CW		
1	T-111	15	2C	224	16	2C52-56	2C44-51	
1	T-204	7	224		27	224		
1	TX-118	6	EB	CW	4	T2-SC	Z/PFP	
1	TY-101	36	1C-F	EB	39	1CFeCN	T1-SC	
1	TY-103	22	TBP	1C-F	9	UR/TBP	T2-SC	
1	TY-104	22	TBP	1C-F	5	UR/TBP		
1	TY-105	37	TBP		5	UR/TBP		
1	TY-106	38	TBP	DIA	5	UR/TBP		
1	U-104	24	R	DIA	23	MW44-51		
1	U-105	6	EB	CW	13	S2-SS	T2-SC	
1	U-107	6	EB	CW	13	S2-SS	MW44-51	
1	U-108	6	EB	CW	13	S1-SC	S2-SS	
1	U-109	6	EB	CW	13	S1-SC	S2-SS	
1	U-111	9	EB	R	13	S2-SS	S1-SC	
1	U-203	18	CW		26	CWR/Al52-60		
Group 2 - Primarily Saltcake Tanks								
2	S-101	1	R	EB	11	R'52-58	S1-SC	High H ₂ O High NO ₃ Low Al High Na
2	S-107	1	R	EB	2	R'52-58	T2-SC	
2	S-109	1	R	EB	20	S1-SC	CWR/Al61-72	
2	S-111	1	R	EB	20	S1-SC	R'52-58	
2	SX-101	1	R	EB	2	R'52-58	S1-SC	
2	SX-102	1	R	EB	20	S1-SC	S2-SS	
2	SX-103	1	R	EB	20	S1-SC	R'52-58	
2	SX-104	1	R	EB	11	S1-SC	R'52-58	
2	SX-105	1	R	EB	20	S1-SC	R'52-58	
2	SX-107	4	R		24	R'52-58	R'59-67	
2	TY-102	2	EB	1C	18	T2-SC	T1-SC	
2	U-102	9	EB	R	40	T1-SC	S2-SS	
2	U-103	9	EB	R	13	S1-SC	S2-SS	
Group 3 - Primarily Sludge Tanks								
3	B-101	6	EB	CW	31	B-SC	BL	Low NO ₃ Low/Medium H ₂ O Low Na High Al
3	BX-101	5	TBP	CW	32	MW44-51	BL	
3	BX-112	8	1C	EB	7	1C44-51/CW	B-SC	
3	C-105	32	TBP	SR-WASH	6	CWP/Al56-60	CWP/Al61-72	
3	C-109	13	TBP-F	1C	8	TFeCN	1C44-51/CW	
3	S-104	4	R		14	R'52-58	CWR/Al61-72	
3	S-110	1	R	EB	11	S1-SC	R'52-58	
3	T-102	17	CW	MIX	6	CWP/Al61-72	MW44-51	
3	T-106	10	1C	CW	38	1C52-56/CW	CWR/Al52-60	
3	U-110	10	1C	CW	10	1C52-56/CW	MW44-51	
3	U-204	39	R	2C	26	CWR/Al52-60		

Results from this cluster analysis were included because they show a few of the broader or more global characteristics of the whole population of single-shell tanks. These results are stated simply as:

1. Group 2 is made up almost exclusively of saltcake waste tanks.
2. Group 3 is made up almost exclusively of sludge waste tanks.
3. Group 1 includes both sludge and saltcake tanks.

This cluster analysis provides a high level check of the agreement between the process history models and the global analytical dataset. From this high level prospective, the process history models are in general agreement with the grouping patterns seen in the analytical data. The process history models were generally able to identify differences between saltcakes and sludges when the analytical data groupings showed separation between saltcakes and sludges (Groups 2 and 3). This may seem like a trivial result. However, it is a good example of a case where there is agreement between analytical data and the process history models.

The results of these global cluster analyses identify some patterns in the analytical data that agree with groupings from both process history models. However, the agreement between the analytical data and the historical models is not strong enough at present. Reliable data on several key tanks is needed before the relationship between analytical data and historical models can be quantified.

4 Between and Within Cluster Similarities

This section discusses the investigation of between-groups and within-groups similarities⁴ on a constituent by constituent basis. This investigation is based on the results from two analysis of variance (ANOVA) models, fit for several constituents considered in the cluster analyses described in the last section.

This section first lists our initial expectations as to the results of an investigation of this type, followed by a description of the ANOVA models and the results from the ANOVAs. Finally, the ANOVA results are discussed.

The following intuitive results are expected from this investigation:

1. The variability between groups is expected to be larger than the tank-to-tank variability within groups.
2. The variability between tanks within a group is expected to be greater than the variability between cores (spatial variability) within tanks.
3. The core-to-core spatial variability calculated using data for tanks across multiple groups should be larger than spatial variability calculated using data from tanks within a single tank group.

The analytical data used in this investigation is a subset of the data reported in the TCRs. For the cations and radionuclides, the analytical results from waste samples prepared by fusion were used. For the anions, water digestion sample results were used.

The first ANOVA model considered is a cross-group model that includes data from all tanks considered in the TCR mean concentration cluster analysis, with the exception of Tank U-110. The analytical data for U-110 was not readily available. This model is expressed as follows:

$$Y_{ijkl} = \mu + G_i + T_{ij} + S_{ijk} + E_{ijkl} \quad (1)$$

where:

Y_{ijkl} is the measured value of an analyte concentration for the l th duplicate from the k th core in Tank j of Group i .

μ is the grand mean concentration of the constituent (averaged across all groups, tanks, cores, and duplicates)

G_i is the deviation of Group i from the grand mean (there are 10 groups)

T_{ij} is the deviation of Tank j within Group i from the grand mean

S_{ijk} is the deviation of Core k within Tank j of Group i from the grand mean (general spatial variability term for all tanks)

E_{ijkl} represents sampling and analytical deviations.

Table 10 contains a summary of the results from the ANOVA run of this model. The table presents estimates for the overall mean, group-to-group variability, tank-to-tank variability within group, core-to-core variability within tank and group, analytical variability, and the number of tanks involved in the analysis for each constituent. The variance component estimates are expressed in the table as relative standard deviations (RSD).

The TCR mean concentration cluster analysis in Section 3.2 revealed a group of five tanks (i.e., Tanks B-110, B-111, BX-107, C-110, and T-104) that are primarily composed of 1C44-51, 1C52-56, and 2C44-51 waste based on the HTCE Model. To investigate this further, a "within-group" ANOVA model was fit to this group, using analytical data from these five tanks.

This within-group model is expressed as follows:

$$Y_{jkl} = \mu + T_j + S_{jk} + E_{jkl} \quad (2)$$

where:

⁴Groups are identified from the TCR mean concentration cluster analysis in Section 3.2.

Table 10: ANOVA Results for 14 Tanks in 10 Groups (Across Groups)

Analyte	Mean	RSD(Mean)	RSD(Group)	RSD(Tank)	RSD(Core)	RSD(Anal)	# Tanks
ICP.f.Al	69800	45	134	0	31	11	14
ICP.f.Bi	43800	38	65	41	11	9	5
ICP.f.Cr	1260	29	77	43	6	15	14
GEA.f.Cs137	38.4	79	0	176	16	12	5
IC.w.F	3520	44	87	6	0	9	5
ICP.f.Fe	18400	17	0	61	21	16	14
ICP.f.Mn	5040	52	152	0	45	20	14
ICP.f.Na	83100	18	19	60	12	8	14
IC.w.NO2	18000	27	0	95	29	15	13
IC.w.NO3	73200	24	47	59	21	11	14
IC.w.PO4	24900	53	147	2	37	15	13
ICP.f.Si	8080	22	13	70	57	15	14
ICP.f.U	25000	82	141	0	0	48	4

Y_{jkl} is the measured value of an analyte concentration for the l th duplicate from the k th core in Tank j

μ is the grand mean concentration of the constituent

T_j is the deviation of Tank j from the grand mean

S_{jk} is the deviation of Core k within Tank j from the grand mean (general spatial variability term for tanks within 1C/2C waste group)

E_{jkl} represents sampling and analytical deviations.

A summary of the ANOVA results is given in Table 11, and discussion follows.

Table 11: ANOVA Results for 1C/2C Waste Type Group (Within Group)

	Mean	RSD(Mean)	RSD(Tank)	RSD(Core)	RSD(Resid)	# Tanks
ICP.f.Al	9.67e+03	32.25	71.63	12.80	7.35	5
ICP.f.Bi	3.50e+04	37.08	51.48	17.16	8.89	2
ICP.f.Cr	1.08e+03	22.57	50.24	6.20	7.95	5
GEA.f.Cs137	8.63e+01	83.29	117.68	8.43	4.37	2
IC.w.F	1.75e+03	8.70	9.79	13.87	3.51	2
ICP.f.Fe	1.85e+04	35.37	78.19	17.10	16.24	5
ICP.f.Mn	1.07e+02	32.81	70.10	28.16	39.51	5
ICP.f.Na	1.15e+04	26.22	58.37	7.71	8.41	5
IC.w.NO2	1.49e+04	50.92	113.10	19.76	13.47	5
IC.w.NO3	1.14e+05	18.96	41.07	16.68	3.45	5
IC.w.PO4	2.23e+04	10.96	19.28	24.00	11.38	5
ICP.f.Si	1.07e+04	28.48	63.19	11.62	8.84	5
ICP.f.U	4.26e+03	33.19	0.00	97.34	44.76	2

Group-to-group variability vs. tank-to-tank variability: For 7 out of 13 constituents, the group-to-group variability is larger than the tank-to-tank variability within groups. This can be seen by comparing the RSDs for the Group and Tank terms in Table 10. Most of the constituents for which this is true had a

large influence on how the tanks were clustered in Table 5. This is an indication that the cluster analysis was effective in maximizing the variability between groups and minimizing the variability within groups (i.e., between tanks within a group). Sodium and nitrate are two of the constituents that had larger tank-to-tank variability than group-to-group variability. This may be due to the soluble nature of these two constituents. Also note that sodium was not a major contributor to group separation in the cluster analysis results in Table 5.

Tank-to-tank variability vs. spatial variability: The tank-to-tank variability is generally larger than the spatial (core-to-core) variability for both the between-group model (Table 10) and the within-group model (Table 11). For the tanks considered, this means that the waste is generally more heterogeneous between tanks than it is within tanks.

Between-group vs. within-group spatial variability: The spatial variability estimated from the between-groups model is generally larger than the spatial variability estimated from the within-groups model. This comparison was made using the raw variance estimates, rather than the RSDs for the core terms. These variance estimates were calculated from the RSDs shown in Tables 10 and 11.

Some of the RSDs in Tables 10 and 11 are reported as zero. This means that the variance contributions of those terms (i.e., terms with zero RSD) are overwhelmed by the variance contributions of the other terms in the model.

5 Pooled Estimates of Spatial Variability

This section demonstrates how to use the cluster analysis results and the estimates of spatial variability in tank characterization efforts.

Several pertinent results from Section 3 impact the analysis and discussion in this section. First, the cluster analysis on the mean concentrations using analytical data, cut at 10 clusters, is used to identify a group of like tanks. An example of the results is presented in Table 5. Second, the results from the cluster analysis on the HTCE estimates will provide an historical perspective on which tanks might be included in this group of like tanks. An example of these results is presented in Table 3. Third, the cluster analysis on the spatial variability, cut at 10 clusters, can be used to assess the appropriateness of pooling estimates of spatial variability across tanks within this group of like tanks. Example results are presented in Table 7. For purposes of this discussion, the pertinent groups of tanks from each of these analyses are presented in Table 12.

Table 12: Cluster Analysis Results from Section 3

HTCE Cluster	HTCE Tanks (from Table 3)	Mean Conc. Tanks (from Table 5)	Spatial Var. Tanks (from Table 7)
7	B-104	B-110	B-110
7	B-107	B-111	B-111
7	BX-107*	BX-107	BX-107
7	BX-108	C-110	T-104
7	BX-110	T-104	
7	BX-112		
7	C-107		
7	C-110*		
7	T-104*		
7	T-107*		
7	TX-109		
19	B-110*		
19	B-111*		
19	T-105*		
19	T-110		

All of the tanks in Table 12 primarily contain 1C44-51, 1C52-56, and/or 2C44-51 waste types. Tanks that were core sampled are identified by an asterisk in Table 12. Tanks that are common to all categories are presented in bold face in the table. Ideally, one would want to use the data for tanks which have been core sampled to make inferences about the tanks in HTCE Clusters 7 and 19 which have not already been sampled. Note that Tanks T-107 and T-105, included in Clusters 7 and 19 from the HTCE analysis, were not included in either the TCR mean concentration or spatial variability cluster analysis groups. Tank T-107 may not group with the other tanks because of sampling and laboratory analytical problems which make the analytical results suspect (see Reference [11]). On the other hand, there is no evidence as to why tank T-105 does not group with the other tanks. Also, although Tank C-110 is contained in the groups for the HTCE and the TCR mean cluster analyses, it did not group with those tanks when the analysis was based on spatial variability. Even though these tanks (T-105, T-107, and C-110) do not follow the ideal expectations and may obscure the overall results, this relationship between tanks is promising and strong.

To use this relationship to make inferences about the tanks that have not yet been sampled, the following approaches are possible:

1. Pool the estimates of spatial variability for Tanks B-110, B-111, BX-107, and T-104. One simple method of pooling variability is to sum the sum of squares for each tank (the variance times the

degrees of freedom⁵) and divide by the sum of the degrees of freedom. Another estimate of pooled spatial variability can be obtained using the estimates of core-to-core variability from ANOVA model results similar to those found in Table 11. An example of pooled spatial variability estimates using the simple method is provided below.

2. If a bounding estimate of spatial variability is desired, then the maximum spatial variability estimate from the tanks that group together based on TCR mean concentrations can be used.
3. For another type of bounding estimate, combine the pooled estimate of spatial variability (from method 1 above) with the tank-to-tank estimate. This bounding estimate relies on model results rather than on extreme value statistics (e.g., maxima); and may or may not be more reliable, depending on the degrees of freedom and the data balance. The bounding estimate becomes more reliable if the degrees of freedom increase and/or the data becomes more balanced.

Table 13 contains an example of pooled spatial variability estimates for a few constituents, computed using the simple method described above. The individual spatial variability estimates for Tanks B-110, B-111, BX-107 and T-104 were pooled in this calculation. The individual tank spatial variability estimates are also shown in the table, for comparison. The pooled estimates are within the range of the individual estimates for each constituent.

Table 13: Pooled and Tank Spatial Variability Comparisons

Tank	Core Samples	Constituent Spatial Variability (Expressed as Standard Deviation)								
		Al	Cr	Fe	Mn	NO2	NO3	Na	PO4	Si
B-110	7	468	33.80	1350	19.10	1040	1390	3770	0	397
B-111	2	313	34.50	1240	0.00	5850	9020	1910	956	1250
T-104	2	2660	0.00	837	0.00	0	0	1620	1190	173
BX-107	2	1430	67.50	1160	7.64	0	12400	369	515	391
Pooled	13	1082	37.42	1270	15.80	2127	5236	3192	537	547

If the difficulties relating to Tanks T-105 and C-110 are neglected, these pooled estimates could be used as spatial variability estimates for other sampled and unsampled tanks in HTCE Groups 7 and 19 (Table 12).

Pooled spatial variability estimates contain more sampling information than the individual tank spatial variability estimates. Due to this fact alone, a 95% confidence interval for the mean concentration could be roughly 80% narrower using the pooled variance, rather than a variance from a tank for which two core samples were taken⁶. This means the overall uncertainty of the mean estimate is reduced when the pooled spatial variability estimate is used.

As stated at the beginning of this section, the results presented here are for demonstration purposes. As more and more meaningful groups of similar tanks are identified, analytical results from these tanks can be pooled together, using methods discussed in this section. These pooled estimates may then be used to help establish inventory estimates (with the associated uncertainties) for tanks where little or no sampling has been conducted; thus reducing the total sampling requirements.

⁵Degrees of freedom in this instance are simply one less than the number of cores taken from a tank.

⁶This calculation assumes that the individual and pooled spatial variabilities are roughly equal. If the pooled estimate is smaller than the individual estimate, the percentage becomes larger. If the pooled estimate is larger than the individual estimate, the percentage becomes smaller.

6 Conclusions and Suggestions for Further Work

The original objectives of this effort were to conduct a comparison between the HTCE and SORWT process history models, to perform an evaluation of the HTCE and SORWT models using analytical data, and lastly, to demonstrate how pooled spatial variability estimates can be used to help reduce sampling requirements. The work in FY95 produced the following results:

1. Based on the analytical data, tanks containing 1C and 2C44-51 (see List of Terms) wastes generally grouped together. This supports suggestions put forth by LANL and WHC scientists that these two waste types are similar, despite the fact that the historical models generate a distinction between them.
2. The available analytical data was used to evaluate the tank groupings based on the two process history models. At present, there is not enough analytical information to perform a complete evaluation of the historical groupings. However, the historical groupings were consistent with some general grouping patterns in the analytical data. For example, where the analytical data showed clear separation between salt cake and sludge tanks, historical groupings showed the same pattern.
3. An investigation was conducted to determine if the list of Hanford Defined Wastes could be reduced from 50 to some smaller number (one suggestion was that there were really only 8 to 10 distinct Hanford waste types). The results of the cluster analysis indicate that 23 of the 50 waste types are distinct. The remaining 27 waste types group mainly as pairs. For example, saltcakes generated from B campaign and from T1 campaign are similar. This result may allow for a small reduction (i.e., from 50 waste types to 36 waste types) in the number of Hanford Defined Wastes; but not the large reduction that has been suggested.
4. The SORWT and HTCE models were compared by performing a cluster analysis on the HTCE estimates and then comparing the tank groups based on the two models. There were some reasonable similarities between the models on a tank-by-tank comparison. On a group-by-group level, the comparison was not as reasonable. Only 7 of the 24 (25%) SORWT groups⁷ remained intact after the HTCE cluster analysis was run. There were only 37% of the single shell tanks that grouped similarly based on the two models. However, using Figure 1 other intersections between the two grouping methods become apparent (i.e., SORWT Group 1 and HTCE Clusters 4 and 20). This information can be used to guide future sampling efforts and data interpretation.
5. A demonstration is presented showing how combined estimates (i.e., combined across tanks) of spatial variability can be calculated. As an example, pooled spatial variability estimates were calculated combining analytical results from Tanks B-110, B-111, BX-107 and T-104. These tanks were selected because they grouped together based on their TCR mean concentrations and spatial variabilities.

This report should be viewed as a demonstration of how a more complete tank grouping study should be conducted when the analytical data becomes available. Core sampling results were available for only 15 of the tanks considered in FY95. The remaining tanks included in this year's efforts used Braun data, which is a less reliable source than core sampling. This effort will continue to be funded in FY96. It is expected that approximately 25 additional tanks will be core sampled and added to the study in FY96. With the addition of these tanks and others in the years to come, more well-defined tank grouping patterns will be identified. These tank grouping patterns can help in allocating future sampling resources more efficiently.

An updated comparison between the HTCE model estimates and analytical results will also be possible in FY96, taking into account the refinements that Los Alamos National Laboratory is making to the HTCE model.

⁷There are 40 SORWT groups but only 24 of them have more than one tank.

References

- [1] Agnew SF, 1994. "Hanford Defined Wastes: Chemical and Radionuclide Compositions," LAUR-94-2657, Los Alamos National Laboratories, Los Alamos, NM.
- [2] Anderson, JD, "A History of the 200 Area Tank Farms," WHC-MR-0132, Westinghouse Hanford Company, Richland, WA.
- [3] Brevick CH, Gaddis LA, Pickett WW, 1994. "Historical Tank Content Estimate for Northeast Quadrant of the Hanford 200 East Areas," WHC-SD-WM-ER-349, Westinghouse Hanford Company, Richland WA.
- [4] Dempster AP, Laird NM, Rubin DB, 1977. "Maximum Likelihood for Incomplete Data via the EM Algorithm," Journal of the Royal Statistical Society, Vol. 39, No. 1, pp. 1-38.
- [5] Everitt BS, 1993. Cluster Analysis, John Wiley & Sons Inc., New York.
- [6] Hill JG, Anderson GS, Simpson BC, 1995. "The Sort on Radioactive Waste Type Model: A Method to Sort Single-Shell Tanks into Characteristic Groups," PNL-9814, Rev 2, Pacific Northwest Laboratory, Richland WA.
- [7] Jungfleisch FM and Simpson BC, 1993. "A Preliminary Estimation of the Waste Inventories in Hanford Tanks Through 1980," SD-WM-TI-057, Rev. 0-A, Westinghouse Hanford Company, Richland WA.
- [8] Little RJA, Rubin DB, 1987. Statistical Analysis with Missing Data, John Wiley & Sons, New York.
- [9] Rencher, AC, 1995. Methods of Multivariate Analysis, John Wiley & Sons Inc., New York.
- [10] Titterington DM, Jiang JM, 1983. "Recursive estimation procedures for missing data problems," Biometrika, Volume 70, No. 3, pp. 613-624.
- [11] Valenzuela BD, Jensen L, 1994. "Tank Characterization Report for Single-Shell Tank 241-T-107," WHC-SD-WM-ER-382, Westinghouse Hanford Company, Richland WA.

APPENDIX A
DATA DESCRIPTION

A Data Description

This appendix explains in detail how a reasonable tank grouping dataset was organized using TCR estimates, TWINS data, and the Braun dataset.

A.1 TCR Estimates

As noted in Section 2 the Tank Characterization Report (TCR) estimates were included in the tank grouping dataset as a first choice.

For the elemental metals and radionuclides, a decision had to be made as to which analytical method and which sample preparation method should be used. For example, there are three estimates for aluminum (i.e., from fusion, acid, and water sample preparations).

The TCR estimates used in the tank grouping dataset were selected based on the following:

1. The estimates used in the TCR to calculate the inventory estimates (i.e., the preferred method estimates) were used whenever they were available.
2. If the preferred method estimates were not available as determined by the TCR writer, the TCR estimates from fusion-prepared samples were used in the tank grouping dataset (Note: For uranium, results obtained by fusion laser fluorimetry were used, rather than those obtained by fusion ICP).

A.2 TWINS Data

The Tank Waste Information Network System (TWINS) data was used as a second source of tank grouping data, when no TCR estimates were available.

All of the analytical results were retrieved from TWINS for all single-shell tanks (excluding those for which TCR estimates were available) and target analyte combinations. The following editing was performed before this set of data was merged with the TCR data:

1. The analytical results from liquid samples were removed from the TWINS dataset, since the tank grouping study is based on the solid phase of the waste.
2. Some unit conversions were made to put all of the analytical results for each analyte in consistent units.
3. Where there were multiple analytical methods/sample preparations, only the analytical results for fusion-prepared samples were retained.
4. For uranium, only the fusion laser fluorimetry results were retained, and all others (including fusion ICP) were excluded from the dataset.

A.3 Braun Data

The Braun dataset was used for the tank cluster analysis only when no sampling information was available from the TCR and TWINS sources. There were several errors and much missing information in the Braun data. For this reason, it was used only as a last resort. Despite the problems with this dataset, however, it contains valuable sampling information and it makes a valuable contribution to the tank grouping dataset. Some revision and editing was carried out on the Braun dataset to improve the results, as follows:

1. Only the analytical results for the analytes of interest in the tank grouping study were retained.
2. All analytical results were converted to $\mu\text{g/g}$ or $\mu\text{Ci/g}$ units, depending on the analyte.
3. There were several analytical results with unrealistically high concentrations. Table 14 lists the concentration extremes for several of the analytes used in the tank grouping study. Analytical results above these extremes were removed from the dataset.

4. The Braun dataset lists several sample phase types. Table 15 shows how these phase types were divided into the more general phases of "liquid" or "solid". Only the solid sample results were used in the tank grouping study.
5. The ICP water digestion results for solid samples were excluded from the revised Braun dataset. The field indicating the analytical method or the sample preparation method was empty for several results. Because of this, several ICP Water analytical results may still be included in the revised Braun dataset.
6. All analytical results from samples taken before a tank's latest sluicing date were excluded from the revised dataset. Obviously, these samples would not represent the waste that is in the tanks at present.
7. Duplicate analyte names were removed from the data set.

Table 14: Braun Analytical Results Extreme Value Cutoff Points

Analyte	Extreme
Aluminum	200,000 $\mu\text{g/g}$
Iron	100,000 $\mu\text{g/g}$
Cesium-137	10,000 $\mu\text{Ci/g}$
Strontium-90	10,000 $\mu\text{Ci/g}$
Uranium	40,000 $\mu\text{g/g}$
Phosphate	200,000 $\mu\text{g/g}$
Plutonium	200 $\mu\text{Ci/g}$
Nitrate	700,000 $\mu\text{g/g}$

Table 15: Division of Braun Sample Types

General Sample Types	Braun Sample Types
Solid	Crust
	Salts
	Sludge
	SludgeComposite
	Solids
Liquid	Filtrate
	Liquid
	Supernate
	Liquor

A.4 Missing Data

Once the three data sources were merged, several missing values were noted in the matrix of data. Since many statistical analysis methods prohibit missing data, the affected tanks would have to be excluded from the analysis. If this proved necessary, however, much sampling information would be lost in the tank grouping study. Therefore, we chose instead to apply statistical methods to estimate the missing values in the dataset, based on the data patterns in the remaining data. References [4], [10], and [8] discuss a technique called the EM algorithm (for Expectation Maximization) that can be used to estimate these missing values using the structure in the data, without building any additional structure into the data. The EM algorithm is a widely used and accepted method for estimating missing values in similar situations.

The variation of the EM algorithm applied to the tank grouping dataset works as follows:

1. First, the missing values for each constituent are replaced with the mean of that constituent across all tanks that produced real analytical results.
2. For each constituent, a regression equation is then constructed which predicts the concentration of the constituent based on the concentrations of all other constituents. For example, the equation

$$Al = \gamma(Bi, Fe, Na, Cs - 137, Sr - 90, U, PO_4^{-3}, Mn, Cr, NO_2, NO_3, F, H_2O, Si) \quad (3)$$

predicts the concentration of aluminum based upon some function γ of the other constituent concentrations. An equation like the one above is obtained for each constituent in Table 2.

3. Predictions calculated from regression equations 3 are used to replace the estimated missing values.
4. Steps 2 and 3 are repeated until relative convergence is obtained (i.e., there is a relatively small difference between the last two consecutive iterations).
5. The missing value estimates obtained after convergence are used in the tank grouping dataset.

These missing value estimates are plotted (and designated by the symbol "M") along with the real data estimates in Appendix A.5.

A.5 Data Plots

This Appendix contains plots of the tank means used in the tank grouping studies, for each analyte in Table 2. Tank means that are large or small relative to the other tank means are labeled with their corresponding tank number. Tank means representing missing values that were estimated are indicated by the plotting symbol "M".

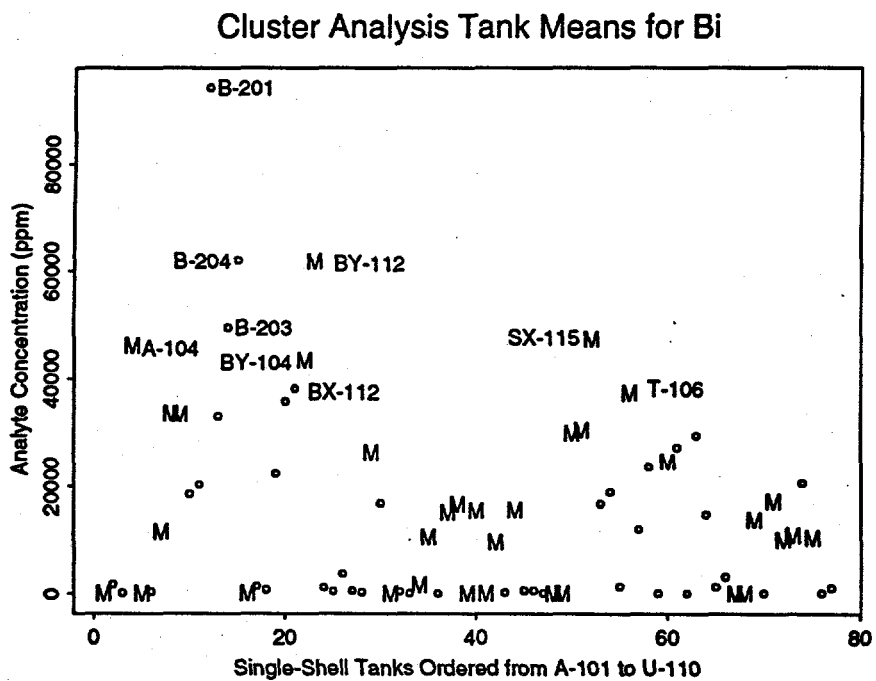
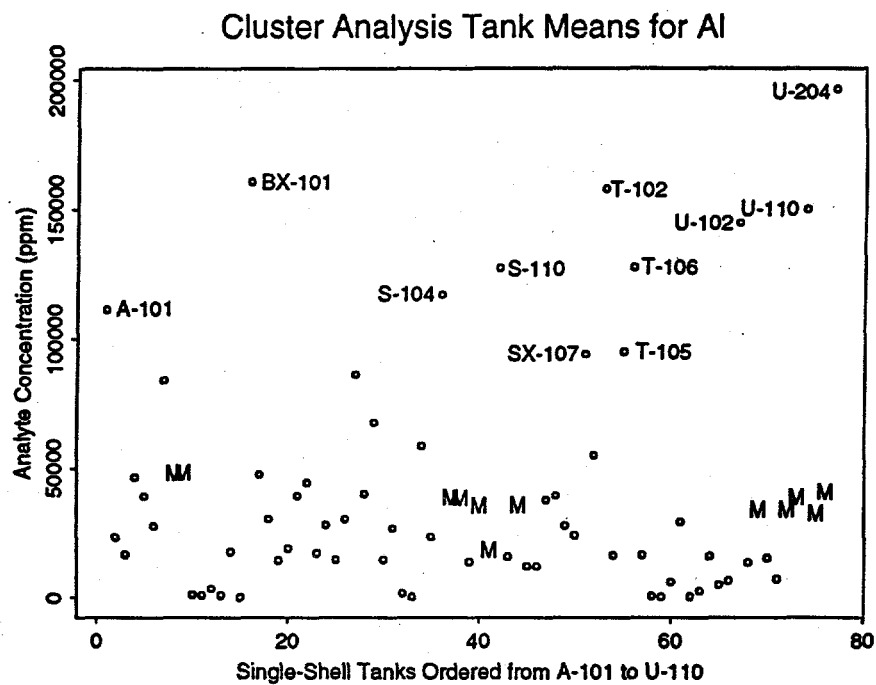
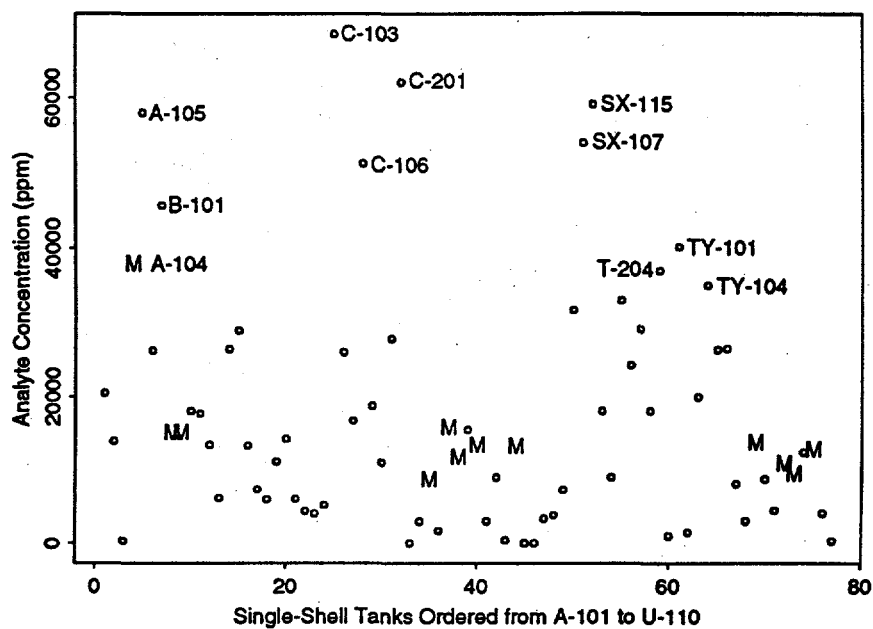


Figure 3: Illustration of Tank Means for Aluminum and Bismuth

Cluster Analysis Tank Means for Fe



Cluster Analysis Tank Means for Na

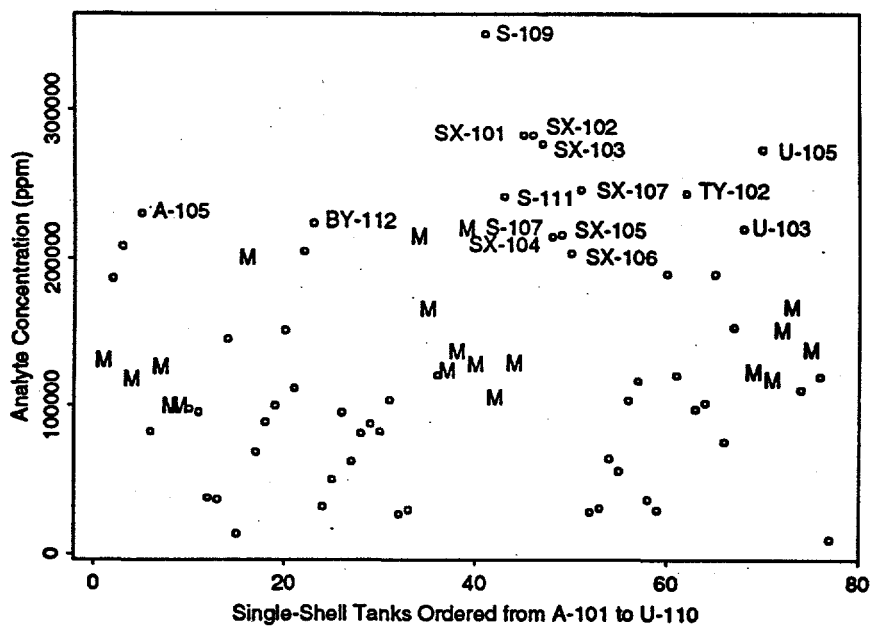
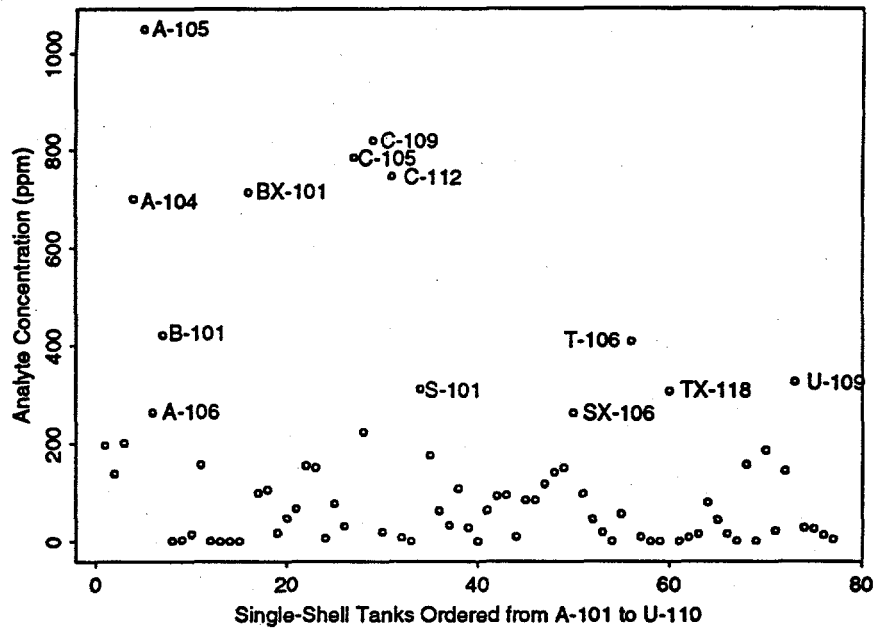


Figure 4: Illustration of Tank Means for Iron and Sodium

Cluster Analysis Tank Means for Cs-137



Cluster Analysis Tank Means for Sr-90

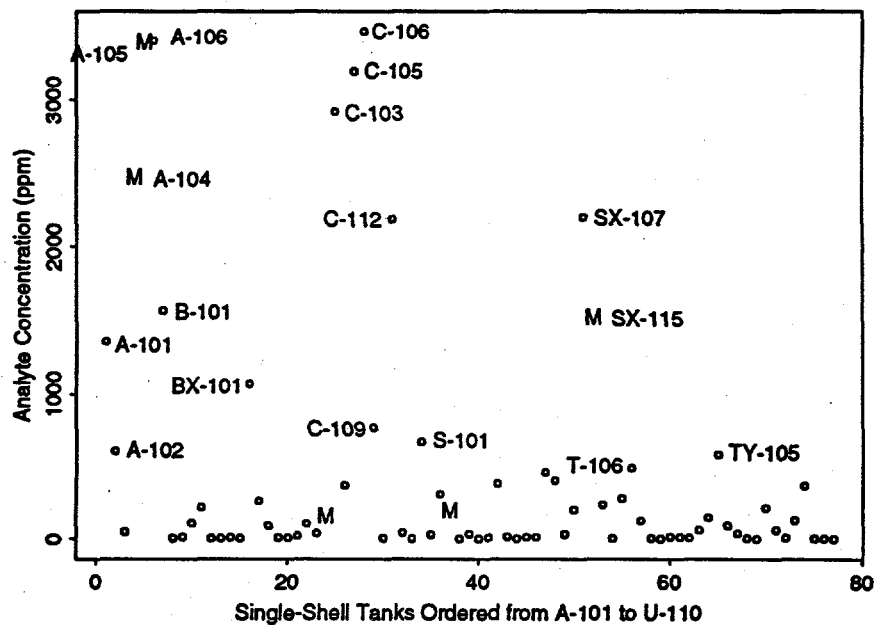
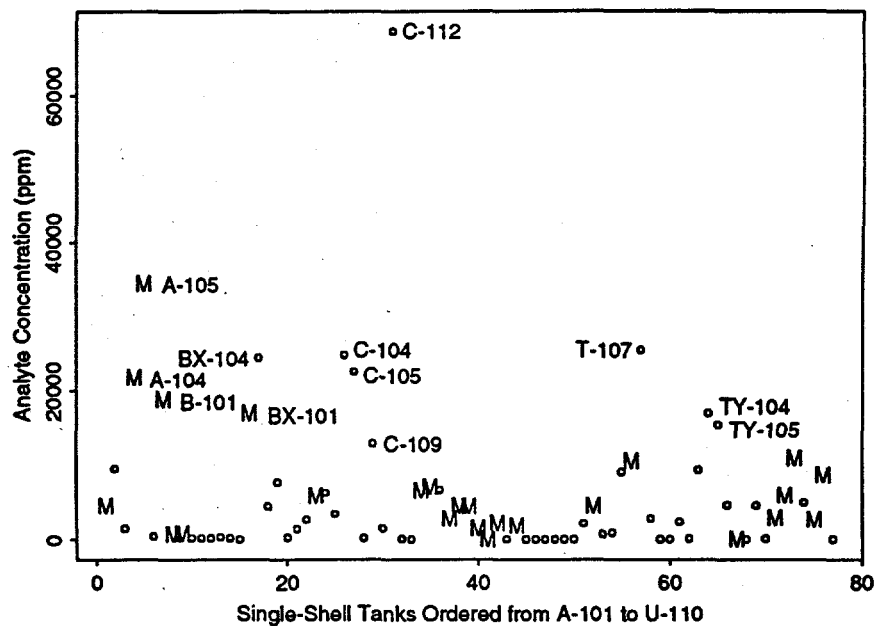


Figure 5: Illustration of Tank Means for Cesium-137 and Strontium-90

Cluster Analysis Tank Means for U



Cluster Analysis Tank Means for Si

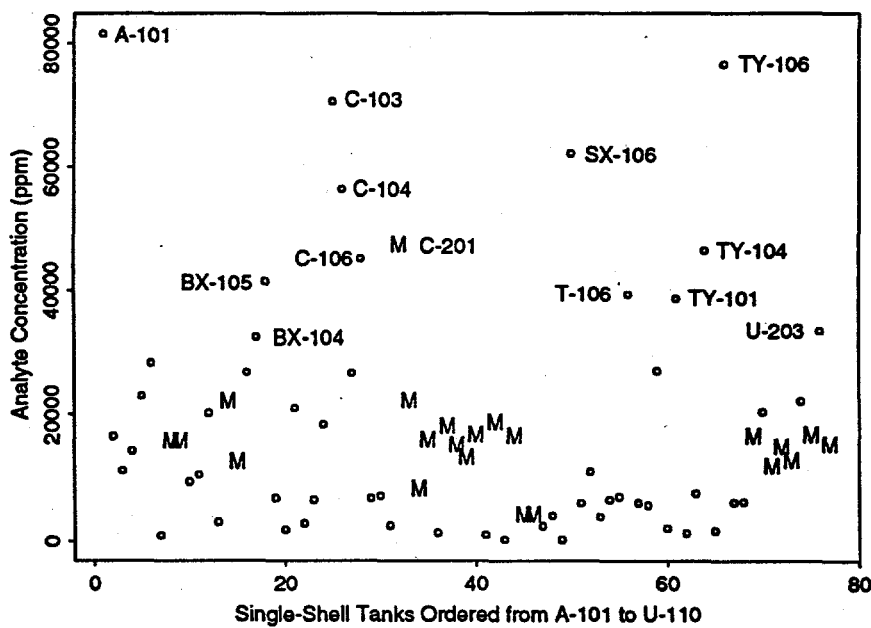
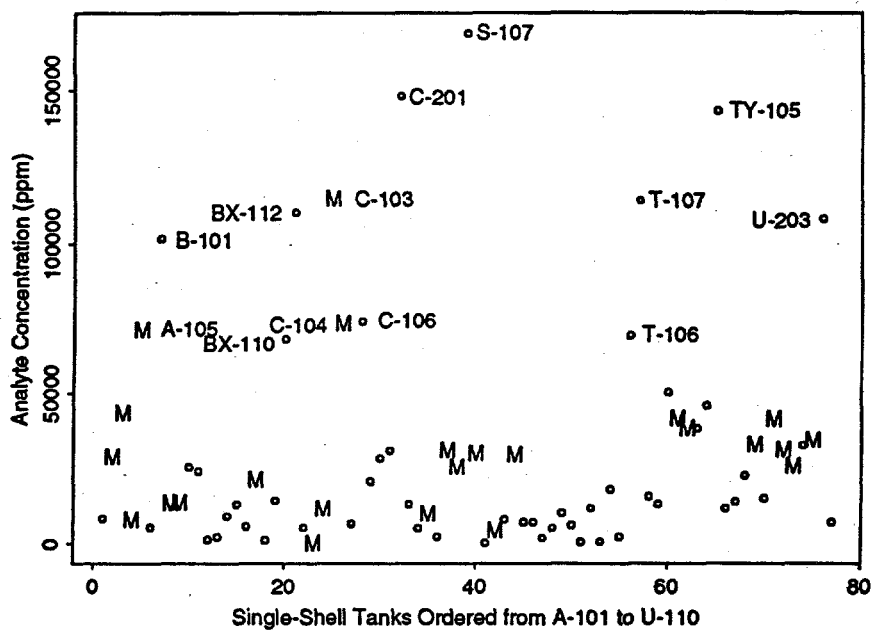


Figure 6: Illustration of Tank Means for Uranium and Silicon

Cluster Analysis Tank Means for PO4



Cluster Analysis Tank Means for Mn

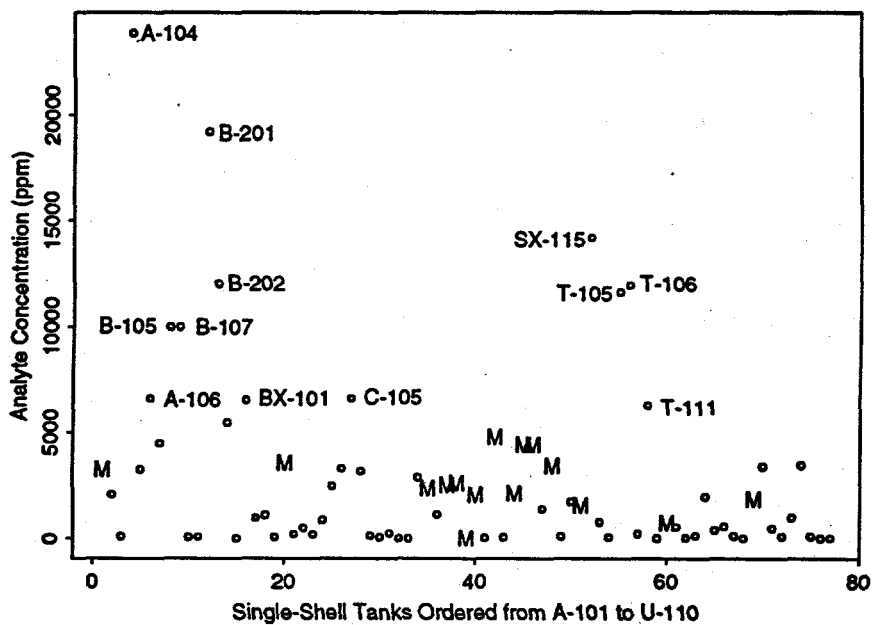
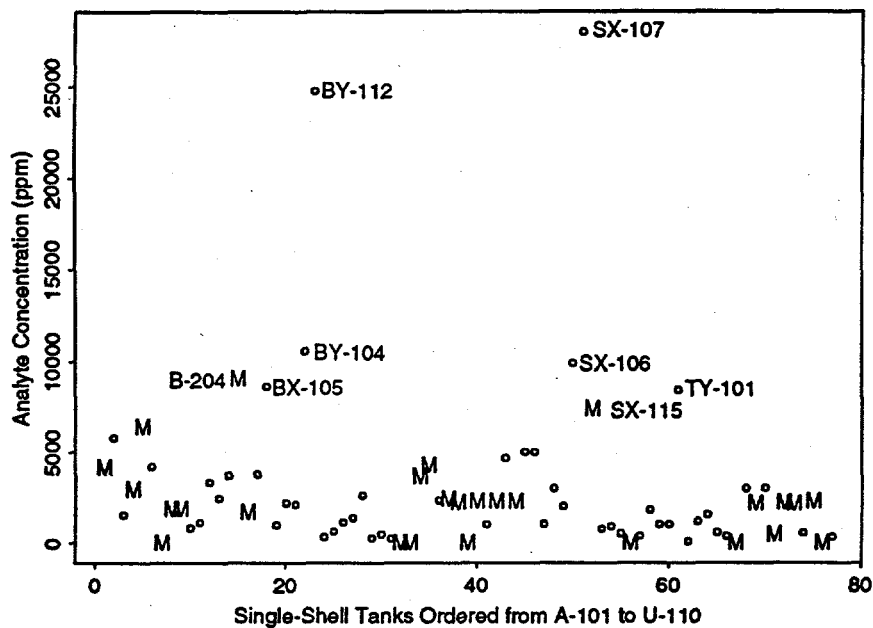


Figure 7: Illustration of Tank Means for Phosphate and Manganese

Cluster Analysis Tank Means for Cr



Cluster Analysis Tank Means for NO2

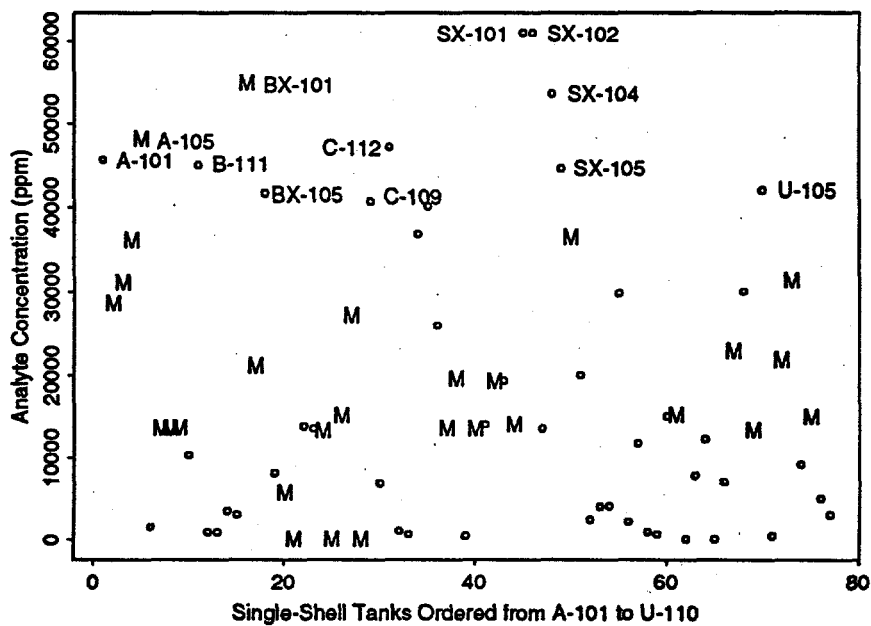
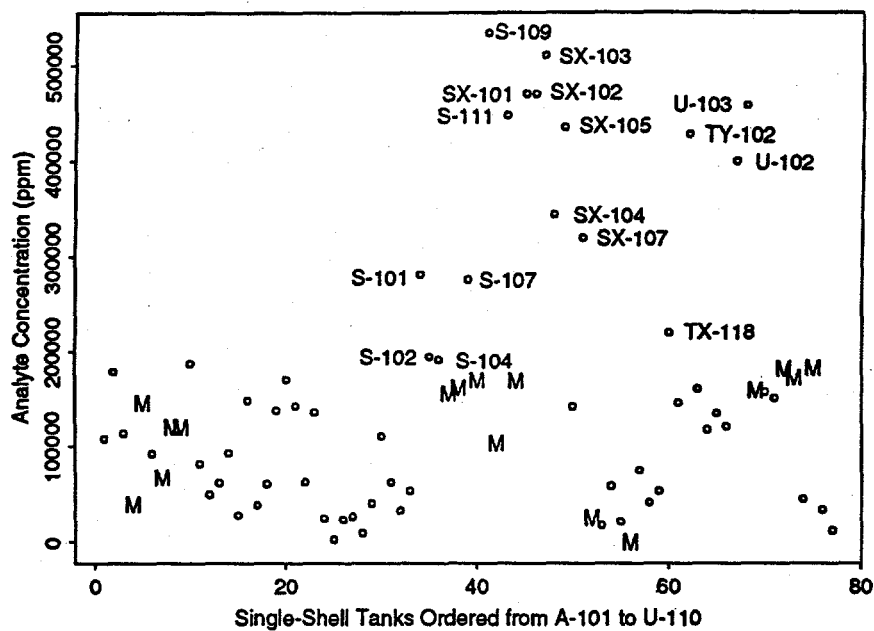


Figure 8: Illustration of Tank Means for Chromium and Nitrite

Cluster Analysis Tank Means for NO₃



Cluster Analysis Tank Means for F

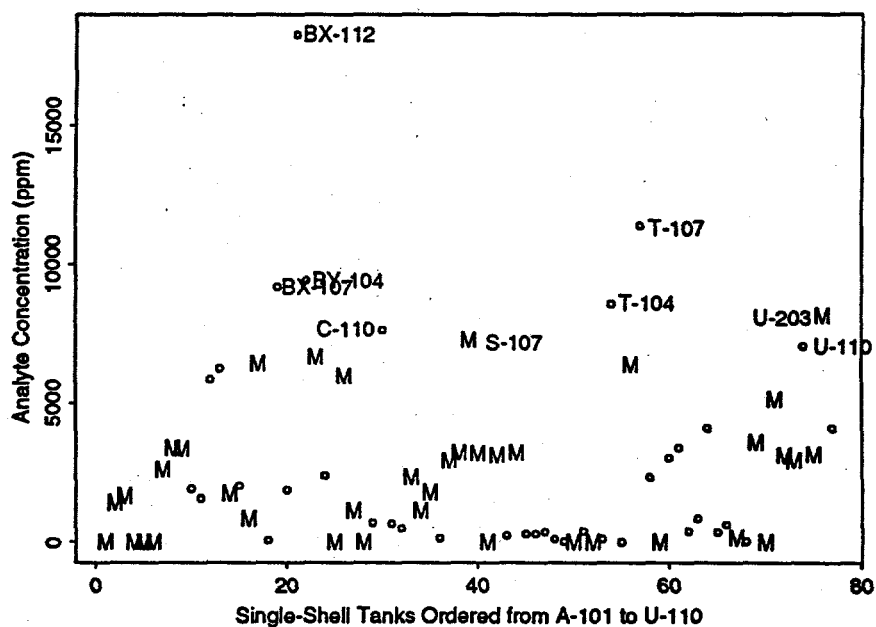


Figure 9: Illustration of Tank Means for Nitrate and Fluoride

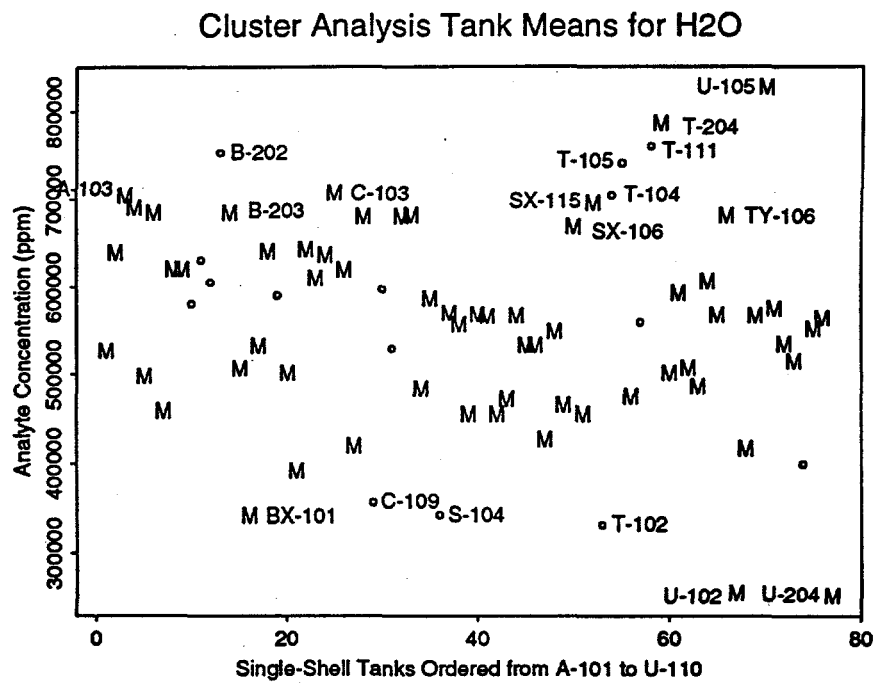


Figure 10: Illustration of Tank Means for Water

DISTRIBUTION

No. of
Copies

No. of
Copies

OFFSITE

2 U.S. Department of Energy
Office of Environmental Restoration and
Waste Management EM-563
12800 Middlebrook Road
Germantown MD 20874
K. T. Lang
J. A. Poppiti

Los Alamos National Laboratories
Group C-346
P. O. Box 1663
Los Alamos NM 87545
S. F. Agnew

SAIC
20300 Century Boulevard
Germantown MD 20874
H. Sutter

ONSITE

3 DOE Richland Operations Office

W. Liou	S7-54
J. F. Thompson	S7-54
N.W. Willis	S7-54

35 Westinghouse Hanford Company

H. Babad	S7-15
R. F. Bacon	S7-85
W. B. Barton	R2-11
T. M. Brown	R2-12
R. J. Cash	S7-15
J. M. Conner	R2-12
D. A. Dodd	T6-50
T. H. Dove	R2-12
S. J. Eberlein	R2-12
L. F. Ermold	S7-84
G. D. Forehand	S7-31
C. S. Homi	R2-12
L. Jensen	T6-07
J. Jo	R2-12
G. D. Johnson	S7-15
N. W. Kirch	R2-11
J. G. Kristofzski	R2-12
E. J. Lipke	S7-14
D. J. McCain	R2-12
J. E. Meacham	S7-15
C. T. Narquis	T6-16
M. A. Payne	S7-84
R. S. Popielarczyk	R1-30
G. F. Raphael	R2-12
L. M. Sasaki	R2-12
B. C. Simpson (2)	R2-12
J. P. Sloughter	H5-27
A. M. Umek	S7-81
D. J. Washenfelder	H5-27
W. I. Winters	T6-50
A. E. Young	T6-06
Central Files	A3-88
TCRC (2)	

No. of
Copies

25 Pacific Northwest Laboratory

C. M. Anderson (2)	K5-12
J. W. Brothers	K5-22
S. A. Hartley	K5-12
P. G. Heasler	K5-12
J. G. Hill	K7-94
L. K. Holton	K9-73
A. M. Liebetrau	K5-12
A. F. Noonan	K9-91
B. A. Pulsipher	K5-12
K. M. Remund (5)	K5-12
W. G. Richmond	K6-51
F. M. Ryan	K7-70
P. D. Whitney	K5-12
Information Release Office (7)	K1-06