DUF6 Depleted Uranium Hexafluoride Conversion Project	**DUF6-G-G-STU-003** **DRAFT FOR UDS REVIEW** **SEPTEMBER 30, 2006** **Cylinder History Study** **Phase II**

Contents Categorization of Paducah DUF_6 Cylinders Using Cylinder History Cards – Phase II

Prepared by:
Henson Technical Projects, LLC
Clinton, Tennessee

Prepared for:
Uranium Disposition Services, LLC
Lexington, Kentucky

Depleted Uranium Hexafluoride Conversion Project

Contents Categorization of Paducah DUF_6 Cylinders Using Cylinder History Cards – Phase II

Lead Preparer

Name: Signature:

Title: Date:

Approval

Name: Signature:

Title: Date:

Approval

Name: Signature:

Title: Date:

Approval

Name: Signature:

Title: Date:

DISCLAIMER

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ACRONYMS AND ABBREVIATIONS

PORTS	Portsmouth Gaseous Diffusion Plant
PAD, PGDP	Paducah Gaseous Diffusion Plant
TRU	transuranic elements
NBIC	National Board of Boiler and Pressure Vessel Inspectors
CID	Cylinder Information Database
NMC&A	Nuclear Materials Control and Accountability
NuMAS	Nuclear Materials Accounting System
UDS	Uranium Disposition Services
DOE	Department of Energy
HTP	Henson Technical Projects
ASTM	American Society for Testing and Materials
ASME	American Society of Mechanical Engineers

Contents Categorization of Paducah DUF_6 Cylinders Using Cylinder History Cards – Phase II

1. BACKGROUND

This Phase II report covers results of the categorization of the contents of individual cylinders for conversion feeding based on retrieval and analyses of the cylinder history cards stored at Paducah.

This Phase II study follows Phase I work[a],which estimated the number of cylinders available which were likely to contain enriched UF6, transuranics (TRU), or technetium (Tc), then individually characterized the specific cylinder population of Model OM cylinders currently stored at Paducah.

The Paducah cylinder cards were maintained over the period from initial large cylinder procurement, until 1988, when a computerized database replaced the card system. The card data traces the usage of cylinders at Paducah (PGDP), but often notes shipment and receipt of cylinders to the Oak Ridge and Portsmouth (PORTS) operating plants.

Previous Phase I work to characterize Model OM cylinders stored at Paducah has been incorporated into this Phase II report as well.

2. SUMMARY OF RESULTS

This Phase II work consisted of analysis of the cylinder cards which traced cylinder history at the Paducah plant. Information collected from the cylinder cards for the cylinders destined for Portsmouth conversion (Oak Ridge and Portsmouth cylinders) was limited to their usage at Paducah. Data for the Portsmouth cylinders will be reported separately.

Summary results of the Paducah Phase II work are illustrated in Figures 1 and 2.

The cylinder card system used by chemical operations at Paducah was replaced by a database in 1988. The database was later transferred to Nuclear Material Control and Accountability (NMC&A). The current database is called NuMAS (Nuclear Materials Accountability System). Analysis of this database, as well as prior accountability records stored on microfiche at Paducah in Phase III should result in the clearing of large numbers of the Model G cylinders that are currently uncategorized. Not including Model G cylinders, over 99% of the cylinder history cards for cylinders at Paducah were located and reviewed.

---

[a] "Historical Preliminary Screening of Uranium Hexafluoride Cylinders for Conversion", Henson Technical Projects, March 2004

Cylinder categories are described below in Section 4. A copy of the procedure used in this categorization is attached as Appendix A. Appendix B consists of the statistical analysis for data validation which was provided by Uranium Disposition Services (UDS) and used to validate these databases. Appendix C is a list of cylinders with rejection history for UF6 processing. Appendix D provides the validation data.

Figure 1. Paducah Summary Data (Number of Cylinders)

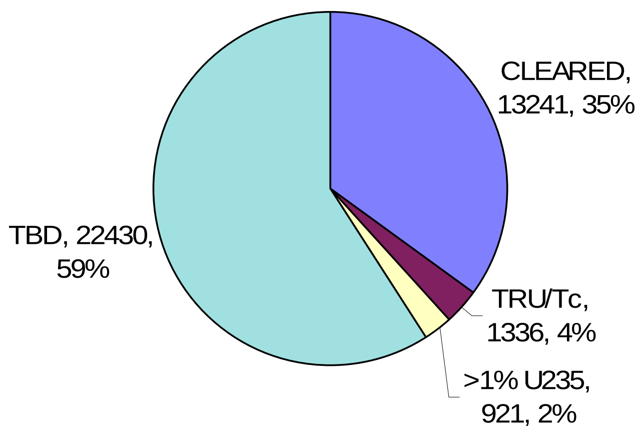
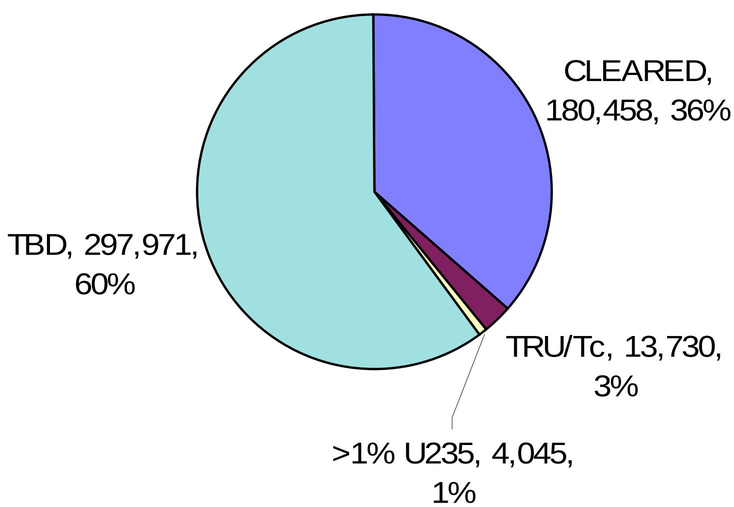


Figure 2. Paducah Summary Data (UF6, in tons)



3. OBJECTIVES, CONSTRAINTS AND ASSUMPTIONS

The objective of this study was to identify, by cylinder number, depleted uranium hexafluoride cylinders containing material enriched to greater than one percent U235, transuranics, and/or technetium (including heel quantities of material), and conversely to clear cylinders containing only uranium hexafluoride.

The primary source used for this categorization study was the set of cylinder history cards which had been maintained at Paducah to track cylinder movement at that site. It was generally assumed that these cards were accurate. However, when process knowledge indicated that the cards might be incomplete, the cylinders were categorized conservatively, so that cylinders were not cleared.

The major constraint of this work was the fact that cylinder history cards were maintained for process tails (depleted UF6) only at Paducah, and that the cards tracked only usage at that plant. The Paducah card system was discontinued in 1988.

It should be noted that all phases of this study are on the cylinder population that existed as of January 2003, prior to movement of any cylinders from Oak Ridge to Portsmouth. This basis allows cylinders to be readily identified with the site where they were last used.

4. CYLINDER CATEGORIES

Cylinders used at the other plants are generally categorized as 4B, (to be determined). As work progresses the number of cylinders in this category will decrease. The categories, with the associated number of Paducah cylinders contained in each category, are summarized in Table 1.

4.1. Category 1 Cylinders

Category 1A cylinders were filled once with depleted tails and placed in storage. There were 9728 Category 1A cylinders documented at Paducah.

Category 1B cylinders were originally filled with either natural feed or depleted tails. This original material was fed or re-fed to the Paducah cascade for enrichment, and the empty cylinder was filled with more tails. These cylinders will not contain appreciable amounts of transuranics or technetium, and should be good candidates for conversion. There were 2681 Category 1B cylinders found.

Category 1C cylinders were used only to hold natural feed or tails after washing and internal inspection. Traditionally, these cylinders were considered clean; however the 832 Category 1C cylinders have been counted separately until any cleaning effectiveness issues are resolved.

4.2. Category 2 Cylinders

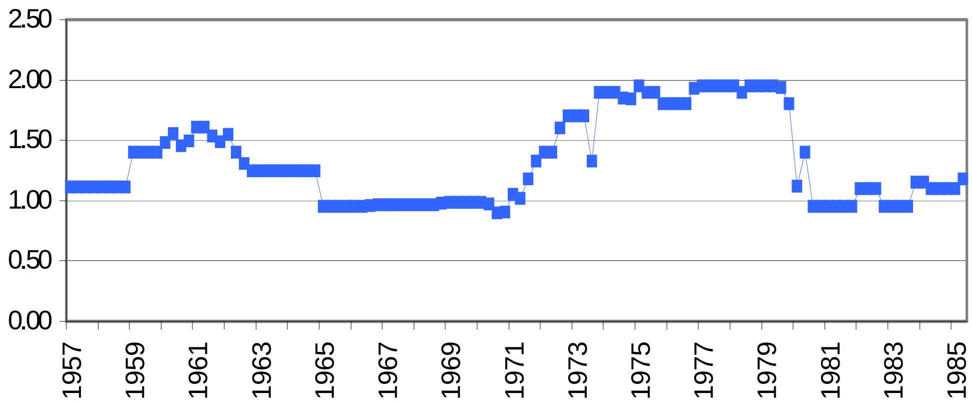
Category 2A cylinders were (generally) filled at the Paducah feed plant, and so might contain residual transuranics and technetium, but never held material enriched to over 1% U235. Cylinders which were used in interplant service are not yet included in this category, but are currently considered to be Category 4B (to be determined). Proposed Phase III work will increase the number of cylinders in this 2A category. Currently there are 1334 Category 2A cylinders.

The sole category 2B cylinder found was filled at least once with Paducah product (at <1% enrichment) which, in the time period these cylinders (other than Model G) were procured and used, contained technetium.

4.3. Category 3 Cylinders

Category 3 cylinders were separated from Category 2 cylinders based on the history of Paducah product U235 assay, as determined by quarterly reports, where available. (see Figure 3).

Figure 3. Paducah Product Assay History



Category 3A cylinders were filled at the Paducah feed plant at least once, and also used to contain Paducah product at over 1% enrichment. There were 584 such cylinders found.

Category 3B cylinders were filled with product (at over 1%) at least once, but never contained recycled material, so would not contain transuranics. If these cylinders were purchased after the Oak Ridge feed plant ceased operation in 1963 (cylinder number 110000, or higher), they were immediately categorized as 3B regardless of interplant service; older cylinders which had been shipped to Oak Ridge are currently contained in Category 4B awaiting further determination of service history. Category 3B presently contains 387 cylinders.

Currently cylinders filled below the top of the Paducah cascade (side withdrawals) are conservatively categorized as Category 3B, if they have not seen feed plant of interplant service. These cylinders are unlikely to contain significant amounts of Tc. The dates of these side withdrawals were noted; and review of the operating logbook records of these cylinder fills are likely to result in clearing of many cylinders for conversion.

4.4. Category 4 Cylinders

Cylinders with missing history cards were categorized as 4A. There were 12,975 such cylinders. Most (approximately 12,000) of these cylinders are Model G cylinders that were filled after the cylinder card system was discontinued.

The 9407 Category 4B cylinders are predominantly those used in interplant service. However a large number of Model O cylinders (almost 5000) with questionable cylinder card history are presently conservatively included in Category 4B, as described below.

The focus of Phase III work at Paducah will be to reduce the number of these cylinders with contents which are dependent on service history at the other two sites.

4.5. Rejected/Category 5 Cylinders

Cylinders which were noted to have been rejected or taken out of UF6 service were categorized separately as R (rejected) in the deliverable database for uploading to Cylinder Information (CID) database, and are described in Appendix C. These rejections are noted independently of CID, and may be used to complement or to confirm data in CID. There were 210 such cylinders identified at Paducah.

Table 1. Categorization of Paducah Cylinders Using Cylinder History Cards

	Category 1: Cleared	Category 2: TRU and/or Tc	Category 3: >1% U^{235}	Category 4: To Be Determined
A	Filled once with natural normal or depleted material. (9728)	Never filled with 1% or greater assay, but have a history of containing recycled feed material. These cylinders may have “hidden heels” containing both transuranics (TRU) and Tc. (1334)	Filled at some time with material >1% assay, and also used to contain recycled material. These cylinders may have “hidden heels” containing both transuranics (TRU) and Tc. (584)	No Paducah history card. (12,975)
B	Filled more than once, but only with natural normal or depleted material. (2681)	No history of recycled feed service, but used to hold Paducah product (at <1% enrichment). These cylinders may also have “hidden heels” which could contain Tc. (1)	No history of recycled feed service, but used to hold Paducah product (at >1% enrichment). These cylinders may also have “hidden heels” which could contain Tc. (387)	History card does not provide enough information. (9407)
C	Washed and subsequently filled with only natural normal or depleted material. (832)		Filled at some time with >1% assay, but have never contained recycled uranium or Paducah product. (n/a for Phase II)	
	CATEGORY 1 TOTAL = 13,240	TOTAL = 1335	TOTAL = 971	TOTAL = 22,382

Table 2. Summary of Results (Number of Cylinders)

CATEGORY	1A	1B	1C	CLEARED	2A	2B	TRU/Tc	3A	3B	>1% U235	4A	4B	TBD	REJECT	TOTAL
MODEL															
30A					303		303	532	137	669	45	772	817	36	**1825**
48A					5		5	17	27	44	7	35	42	2	**93**
48X			2	2	42		42	11	4	15	63	46	109	44	**212**
48T			1	1			0		162	162	5	408	413	29	**605**
48O		121	61	182	248		248		5	5	7	4843	4850	10	**5295**
OM	2111	2341	284	4736	631	1	632	11	2	13	148	1868	2016	31	**7428**
G	7307	219	484	8010	102	0	102	13	0	13	12683	1374	14057	33	**22215**
48H	15			15							1		1	6	**22**
48HX	8			8								7	7	19	**34**
48Y												60	60		**60**
48OHI									0	0		1	1		**1**
12-in					4		4				15	42	57		**61**
CV	287			287									0		**287**
TOTALS	**9728**	**2681**	**832**	**13241**	**1334**	**1**	**1335**	**584**	**387**	**971**	**12975**	**9407**	**22382**	**210**	**38138**

Table 3. Summary of Results (Amount of UF6, in Tons)

CATEGORY	1A	1B	1C	CLEARED	2A	2B	TRU/Tc	3A	3B	>1% U235	4A	4B	TBD	REJECT	TOTAL
MODEL															
30A					684		684	1205	322	1527	85	1780	1865	3	4079
48A					32		32	132	207	339	40	245	285	18	674
48X			21	21	414		414	98	29	127	309	36	345		207
48T			11	11					1,653	1,653	11	3,891	3,902	22	5,588
48O		1,697	853	2,549	3,465		3,465		70	70	70	67,767	67,837	56	73,978
OM	29,510	32,738	3,942	66,189	7,819	14	7,833	153	27	180	1,955	25,738	27,693	68	101,962
G	101,355	3,050	3,050	107,455	1,302		1,302	149		149	176,099	19,088	195,187	111	304,204
48H	111			111							14		14	84	209
48HX	27			27								13	13	67	107
48Y												811	811		811
48OHI												14	14		14
12-in											1	3	5		5
CV	4,095			4,095											4,095
															0
TOTALS	135,097	37,485	7,876	180,458	13,716	14	13,730	1,736	2,309	4,045	178,584	119,388	297,971	428	496,633

5. DATA SUMMARY

Of the 38138 cylinders at Paducah (as of January 2003), 13241 were categorized as Category 1 cylinders; cleared for feeding. These included cylinders with only one fill with depleted UF6, multiple fills with normal enrichment and depleted UF6, and cylinders filled with depleted UF6 after cleaning and internal inspection. These conditions are listed as sub-categories in Table 1. This separation of each category into subcategories would allow for ready re-categorization at Uranium Disposition Services discretion.

Category 2 cylinders could contain residual amounts of transuranics and/or technetium; there were 1335 such cylinders. Category 2a cylinders had been used for service at the Paducah Feed Plant and could contain residual TRU and Tc. Category 2b cylinders had been filled with Paducah product, and could contain residual Tc (but not TRU). Most product cylinders were shipped to the other two sites, so that only one of these cylinders' histories could be completely determined through the Paducah cylinder card history.

Category 3 cylinders may contain heels with enrichments over 1%, as well as heels containing technetium. There were 971 such cylinders.

There are currently 22382 Category 4 cylinders with not enough information on the cylinder history cards for complete characterization. These include a) cylinders with no history cards and b) cylinders with incomplete history card information, such as some cylinders with service at other site(s).

Results of this Phase II work are summarized, by cylinder model in Tables 2 and 3, and are described in more detail below)

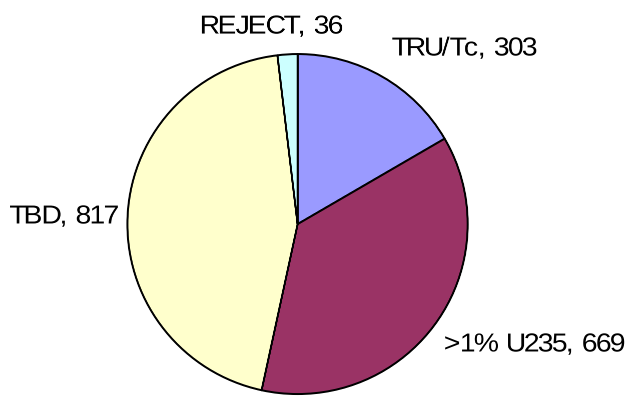
5.1. Model 30A Cylinders

Hundreds of government surplus chlorine cylinders (known as Model 30A cylinders), manufactured by Columbiana Boiler Works, were modified in Oak Ridge for use as UF6 tails cylinders. The 30A cylinders were also used as feed cylinders when the Oak Ridge UF6 feed plant started operations in 1951.

These are the oldest storage cylinders in cylinder yards today. The earlier depleted material stored in the smaller MD cylinders was re-fed to the cascade in concerted efforts to support increased production until the new Oak Ridge K-1131 feed plant was operating at capacity.

The categorization summary for Model 30A cylinders is shown in Figure 4. No 30A cylinders have been cleared to date.

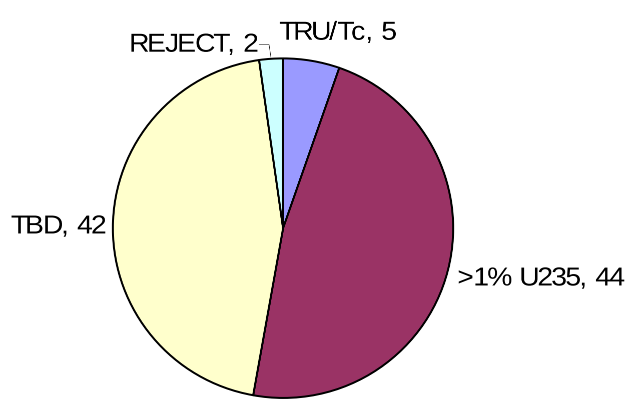
Figure 4. Distribution of Paducah Model 30A Cylinders



5.2. Model A Cylinders

These early non-Code-qualified cylinders, a total of 2865 according to procurement records, were rated for a working pressure of 200 psig and had a (nominal) wall thickness of 5/8-inch. They were fabricated from A285 Grade C pressure vessel plate. Model A cylinders do not have certified volumes. No Model A cylinders were cleared in Phase II.

Figure 5. Distribution of Paducah Model 48A Cylinders



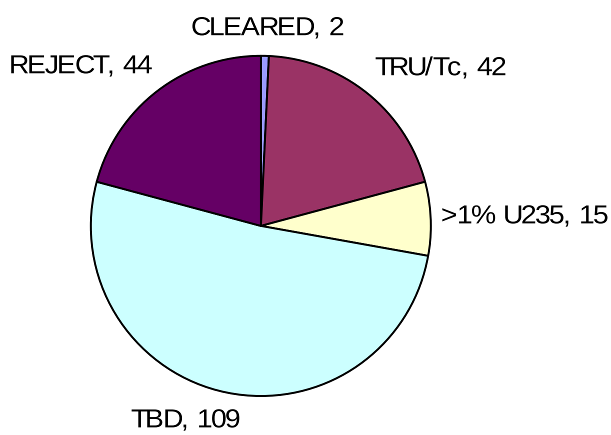
5.3. Model X Cylinders

In 1953 and 1954 a group of 1500 Code-qualified Model X cylinders were procured. The A285 steel cylinders, of 5/8-inch wall thickness (heavy-wall) were used for interplant transport, as well as tails storage. Model X cylinders have certified volumes, which were determined by water weight at the time of manufacture.

In 1989 DOE purchased Model X cylinders from TVA and Sequoyah Fuels surplus stocks. These cylinders were of post-1980 manufacture and were fabricated from ASTM A516 plate. They were Code-qualified and had certified volumes. A total of 356 cylinders were involved in these purchases.

Distribution of the Model X cylinders at Paducah is shown in Figure 6. Only two of 212 cylinders were cleared using cylinder history cards.

Figure 6. Distribution of Paducah Model 48X Cylinders

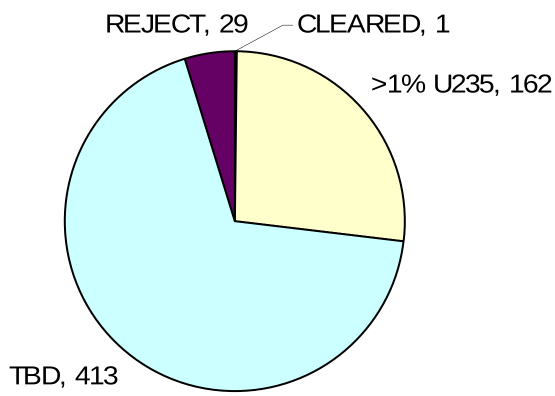


5.4. Model T Cylinders

The Model T cylinders were the earliest 5/16" wall thickness (thin-wall or LSA) cylinders, and both skirts and external stiffening rings were of thinner material than in the earlier heavy-wall design. Procurement records show that 4230 Model 48T 10-ton cylinders were purchased to this design in the 1956-58 time frame. These cylinders were fabricated from A285 steel, and were Code-qualified, but did not have certified volumes.

It should be noted that the Model T cylinders were frequently used as Paducah product cylinders. They were all originally delivered to Paducah; those in Oak Ridge and Portsmouth were shipped there from Paducah. At least some of these cylinders were shipped full to Oak Ridge. There are 605 Model T cylinders at Paducah; their categorization is shown in Figure 7.

Figure 7. Distribution of Paducah Model 48T Cylinders

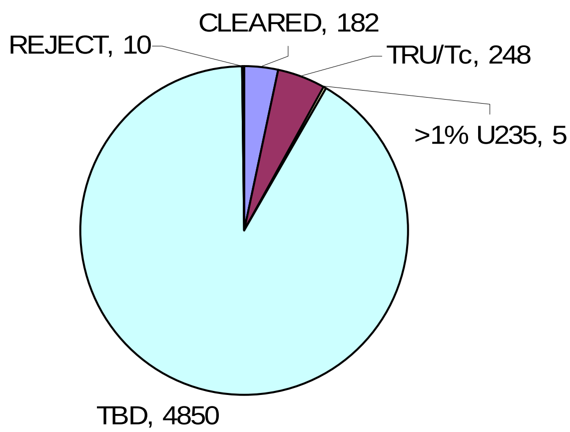


5.5. Model O Cylinders

The 14-ton capacity scale-up of the 48-inch thin-wall cylinder is the Model O cylinder. While maintaining the 48-inch diameter of the 10-ton Model T cylinder, this version gained additional capacity by increasing the body length. The three external, solid stiffening rings were replaced by two 6-inch structural channels. Skirts were omitted on this cylinder design. The Code-qualified cylinders were fabricated from A285 pressure vessel plate. A total of 6602 were purchased between 1958 and 1961; 5295 are at Paducah. Model O cylinders do not have certified volumes.

A large number of the Paducah Model O cylinder cards showed abnormal cylinder flow; with no identification of the tails withdrawal facility, and unusual yard placement for normal tails. This may be an anomaly of history cards from that period. Review of accountability records may clear most of the 4850 of these cylinders that are conservatively not categorized with the available history card information.

Figure 8. Distribution of Paducah Model 48O Cylinders

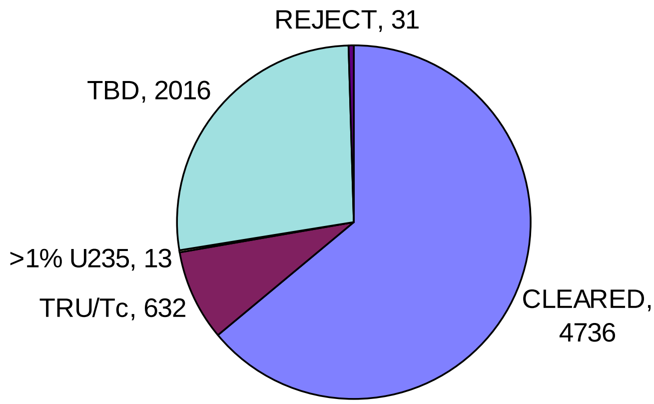


5.6. Model OM Cylinders

The Model OM, a modification of the Model O design, returned to the three solid stiffening rings of the earlier 10-ton design. The Model OM cylinders were fabricated of A285 steel, from 1962 to 1978. The initial version of the Model OM thin-wall 14-ton cylinder had three 11/8-inch thick stiffening rings; stiffening ring specifications were changed to 1 inch thick for later procurements. There are 7428 Model OM cylinders at the Paducah plant. Of these 4365 are considered cleared for conversion, as shown in Figure 9.

The Paducah Model OM cylinders were the initial cylinders individually categorized for the cylinder history study. After the initial work was completed, over 300 additional history cards were found for these cylinders, and the dataset has been modified to reflect these new data.

Figure 9. Distribution of Paducah Model 48OM Cylinders



5.7. Model G Cylinders

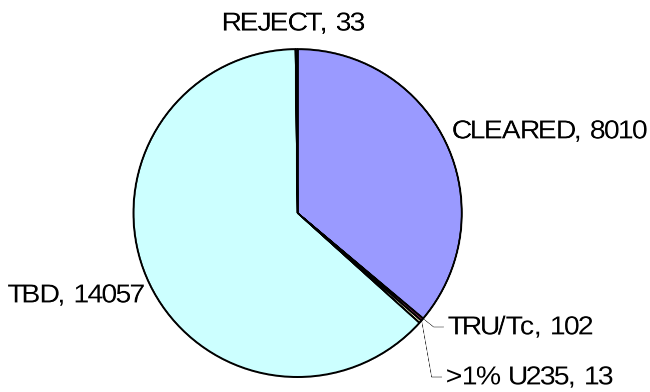
The Model G thin-wall storage cylinder was introduced in 1977, as a redesignation of the Model OM cylinder. The two cylinders are identical in all respects. Model G cylinders are Code-qualified and have certified volumes.

The initial procurement lot of 3000 cylinders was fabricated from A285 plate in 1978 through 1980. All later procurements specified A516 steel, with a modified (low-sulfur) composition specified after 1983 deliveries.

There were 22,215 Model G cylinders stored at Paducah at the initiation of this cylinder history study. Of these, 8010 have been cleared using cylinder history cards. There are 12,683 cylinders without cylinder history cards; almost all of these were filled after the cylinder history card system was abandoned in 1988. These should be readily cleared using uranium accountability records in Phase III. Most of the Model G cylinders were filled after Tc levels in Paducah product fell below ASTM standard limits in 1982.

All Model G cylinders were filled after the Oak Ridge and Paducah feed plants, which were used to produce UF6 from recycled (as well as normal) uranium oxide, were shut down; so they should not, in general, contain TRU material. However, as seen in Figure 10, 115 Model G cylinders have history card noting recycle material fills, or fills with <1% enriched material. These isolated cases were likely the result of immediate cylinder transfer needs, and are expected to represent less than 5% of the Model G population. Some of these cylinders may be full of recycled UF6, although they were filled after both feed plants had been decommissioned.

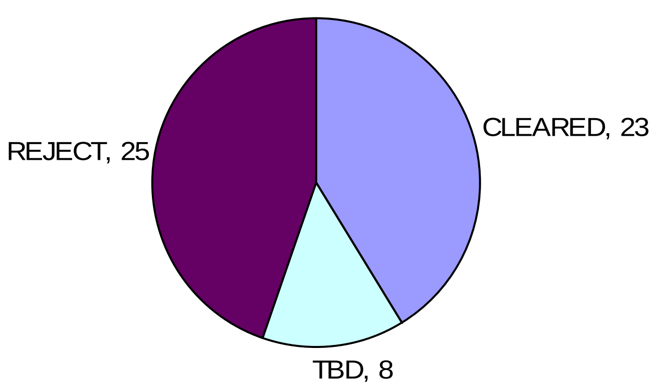
Figure 10. Distribution of Paducah Model 48G Cylinders



5.8. Model H and HX Cylinders

These are dimensionally identical to the Model G cylinder but are fitted with valve and plug end skirts. The cylinders are Code-qualified, and have certified volumes. The HX cylinders, 1000 in number, were fabricated of A285 steel, and were delivered between January and June of 1979. Later Model H procurements totaling about 3100 cylinders between 1979 and 1997 were fabricated of A516 steel. There are 56 of these cylinders at Paducah, with distribution as shown in Figure 11.

Figure 11. Distribution of Paducah Model 48H and HX Cylinders



5.9. Model Y and OHI Cylinders

The heavy-wall Model Y cylinder was skirted at both ends, fabricated of ASTM A516 pressure vessel plate and were Code-qualified. Two procurements in this design totaled 260 cylinders. Model Y cylinders have certified volumes.

These cylinders were heavily used for interplant transport, primarily of tails material. All 60 of the Model Y cylinders at Paducah could not be categorized because of service at the other two sites. The single Model OHI cylinder at Paducah had also seen interplant service.

5.10. Model 12A and 12B Cylinders

Categorization of only four of the 61 12-inch cylinders at Paducah was possible. These cylinders were often used for off-site shipments and as cold traps.

5.11. Model CV Cylinders

In 1958, the Paducah Gaseous Diffusion Plant fabricated 287 cylinders from nickel-lined converter shells specifically for cost-effective storage of depleted UF6. These cylinders were not Code-qualified. (It is believed they were hydrostatically tested to 200 psig to qualify them for their 100-psig design working pressure.)

Engineering drawings are available for the CV cylinders, and some design data. These vessels were fabricated to the same criteria as ASME code vessels including qualified welders and weld inspections. However, they were not fabricated by a Code shop and are not Code-qualified vessels.

Following the procedure detailed in Appendix A, one CV12 cylinder card and 55 CV19 cylinder cards out of a total of 287 containers were judged to not contain enough information to characterize the containers. One additional CV12 card was missing.

However, since there was no equipment available at Paducah for feeding UF6 from these containers, they were all categorized as “single fill with tails”.

6. DATA VALIDATION

Data validation was performed as work progressed according to the two-step procedure outlined in Appendix B. The Paducah cylinders were divided into categories as follows for checking:

- Model G cylinders (with cards)
- Model OM cylinders
- Miscellaneous Paducah 48-inch cylinders
- Paducah Model 30A cylinders
- Paducah 12-inch cylinders

Appropriately sized samples of cylinders for each of these groups were randomly selected, and blank data sheets were furnished for the effort. Cylinder categorization was performed by a second evaluator, who had received previous training in the categorization procedure. Copies of the data validation sample results are included as Appendix D.

APPENDIX A. PROCEDURE FOR CYLINDER CATEGORIZATION FOR CONVERSION PROCESSING, BASED ON CYLINDER HISTORY CARDS

Scope: This procedure is to be used for the categorization of Paducah Uranium Hexafluoride (UF6) Cylinders for feeding to the conversion plant. This categorization is primarily based on cylinder movement information obtained from the Paducah Gaseous Diffusion Plant Cylinder History Cards.[a]

Required Materials:

Cylinder history cards
Copies of Table 2 (electronic or paper) and CID Printout
Excel Database downloaded from CID

Categorization Steps: Cylinders will be divided into three categories based on information such as movement, locations, and weights of the cylinders, as well as process knowledge. A fourth category will be used for cylinders that do not have a Cylinder History Card, and cylinders for which the card does not provide sufficient information to categorize. Table 1 provides the descriptions of these categories. Cylinders with “reject” anomalies noted on the Cylinder History Card will be classed as REJECT. All categorizations will be done conservatively.

Review each cylinder history card and complete Table 2.
Determine category based on the identified cylinder service history noted on Table 2.
Enter cylinder categories in Excel Database.

Disposition of records: Cylinder History Cards: Any Cylinder History Cards retained by HTP, as appropriate and approved, will be maintained as Quality Level 2 records.

One electronic copy of each completed Table 2 will be provided to UDS as back-up information for the cylinder categorization; one copy will be retained by HTP for a period of one year past project termination.
One electronic copy of the EXCEL database will be retained by HTP for a period of one year past project termination; a second copy will be furnished to UDS for external QA review and incorporation into CID.
CID printouts, as annotated with working notes, will be retained by HTP for a period of one year past project termination.

Data Validation: Data validation will be performed according to approved statistical methods, as requested by customer.

---

[a] These cards generally do not have any information on the physical condition of the cylinders/containers. Further categorization based on physical container condition, on definitive Nuclear Material Control and Accountability assay records, or on diffusion plant operating logs would be a separate effort from the work described in this procedure.

Table 1. Cylinder Categories

	Category 1: Cleared	Category 2: TRU and/or Tc	Category 3: >1% U235	Category 4: Unknown
A	Filled once with natural normal or depleted material.	Never filled with 1% or greater assay, but have a history of containing recycled feed material. These cylinders may have "hidden heels" containing both transuranics (TRU) and Tc.	Filled at some time with material >1% assay, and also used to contain recycled material. These cylinders may have "hidden heels" containing both transuranics (TRU) and Tc.	No Paducah history card.
B	Filled more than once, but only with natural normal or depleted material.	No history of recycled feed service, but used to hold Paducah product (at <1% enrichment). These cylinders may also have "hidden heels" which could contain Tc.	No history of recycled feed service, but used to hold Paducah product (at >1% enrichment). These cylinders may also have "hidden heels" which could contain Tc.	History card does not provide enough information.
C	Washed and subsequently filled with only natural normal or depleted material.		Filled at some time with >1% assay, but have never contained recycled uranium or Paducah product. (n/a for Phase I)	

Table 2. Cylinder Categorization Form

("√" in applicable columns)

Category	Sub-Category	Cylinder Number	Tare Weight Date	Feed Plant Service	Product Withdrawal	Single time tails fill	Multiple tails fill with only tails or incoming from PORTS?	Incoming from OR w/ # > 110000	Incoming from OR w/ # < 110000	Any Noted inspection rejection?	Notes (e.g., pertinent dates, reject information, other)

Legend:

Single or multiple tails fill = 1A or B, respectively
Feed Plant Service = 2A or 3A
Product Withdrawal w/o FP= 2Bor 3B
Incoming from OR = 4B
No History Card = 4A
Inspection rejection noted=4C

APPENDIX B: STATISTICS

Instructions and Tables of Sample Sizes for Using Statistical Two Stage Sampling Plans To Verify the Accuracy of the Electronic Data Files on the DUF_6 Cylinders

Date: September 29, 2004

To: Scott Barnes, Duratek Federal Services
Copy: Tom Price, Fairfield Service Group

From: Jack Zolyniak, Statistical Consultant

Subject: Instructions and Tables of Sample Sizes for Using Statistical Two Stage Sampling Plans To Verify the Accuracy of the Electronic Data Files on the DUF_6 Cylinders

Uranium Disposition Services (UDS) is transferring information from historic paper records to electronic data files on approximately 62,000 cylinders of DUF_6 located in Oak Ridge, Paducah and Portsmouth. The electronic data are an integral part of UDS' strategy to sample, process, and dispose of the cylinders.

Below you will find instructions for using statistical two stage sampling plans to verify the accuracy of the electronic data files, and two short tables of sample sizes. The tables can easily be used to determine sample sizes, using only the number of data records to be checked for errors. To assure the sampling plans are applied in an effective and consistent manner, UDS should establish rules for comparing the electronic data with the paper records and for performing appropriate corrective actions, as necessary. I offer some practical suggestions for the two sets of rules. You will also find brief discussions on quality properties of the sampling plans, the graph of sample sizes versus files sizes, construction of the sample size tables, the choice of the sampling plans, and a list of statistical references.

Quality Properties

From our prior conversations, the accuracy of the data files will be verified using statistical two stage sampling plans. Stage 1 and stage 2 sample sizes were calculated so probability statements (1) and (2) are true. For the purposes of the statistical sampling, the statements specify values of *acceptable and unacceptable error rates* for the data files. When taken together, the statements establish *quality performance properties of the sampling plans*.

(1) Data files that have no more than 0.5% errant records will be found *acceptable* by the sampling plans no less than 99.5% of the time.

(2) Data files that have 5.0% or more errant records will be found *unacceptable* by the sampling plans no less than 95.0% of the time.

Data files that have error rates between 0.5 and 5.0 percent will be found acceptable between 99.5 to 5.0 percent of the time by the sampling plans. Files with no errant records will be found acceptable 100% of the time; files with all errant records will never be found acceptable. Note that some error rate will be found acceptable 50% of the time. This is called *indifference value*. All of these are points on the Operating Characteristics graph for the statistical sampling plans. Please see page 8.

Two Stage Plans Can Reduce the Numbers of Samples Checked

Two stage sampling plans can reduce the number of records that are sampled compared to single sample sampling plans with the same properties. This occurs most often for data files that are less than 0.5% deviant or more than 5% deviate when an ⅓cceptable or unacceptable conclusion is reached without taking the second stage sample. The first stage sample is usually smaller than the sample in the single sampling plan. Data files with error rates between 0.5% and 5.0% will frequently require taking the second sample. The sum of the two samples is larger than the sample in the single sample plan.

Operation of the Two Stage Sampling Plans

The sampling plans are specified by two sample sizes (n_1, n_2) and two acceptance numbers: (c_1, c_2). The number of errant records found in the sample(s) is compared to the acceptance numbers to determine if the data file has acceptable or unacceptable quality.

Suppose N data records will be checked for entry errors. Use either Table 1 or Table 2 to find the two sample sizes. Randomly select n_1 data records from the N records and check the data entries for agreement with the paper records. Count the number of errant records in the sample. If there are c_1 or fewer errant records, the N records are found to have acceptable quality. If there are more than c_2 errant records in the sample, the N records are found to have unacceptable quality. If the number of errant records is more than c_1, but c_2 or fewer, the second sample is needed. Randomly select n_2 data records from the group of N-n_1 unchecked records and check the entries for agreement with the paper records. If the *total number* of errant records in the two samples is c_2 or fewer, the N records are found to have acceptable quality; otherwise, the N records are found to have unacceptable quality. Please see page 5 for a data checker's outline of the sampling plans and pages 6 and 7 for Tables 1 and 2.

Note: Historically, these plans have been called *double sampling acceptance plans for attributes*. Here, the attribute of interest is an accurate data record and the number of inaccurate records is determined by *counting* (integer values). There are also sampling acceptance plans *for variables* where a measurement is obtained for each item in a sample.

For example, a diameter is measured, or a chemical concentration level is calculated, etc. The distinction is statistically important.

Options with Data Files

The sampling plan may be applied to all or part of a data file. The sampling plan may also be used on a group of N data records before the entire data file has been completed. The conclusion of the sampling plan applies only to the N records that are sampled. Tables 1 and 2 show that the sampling plans are most efficient when a large number of records are sampled.

Corrective Actions

UDS should establish the corrective actions that will be taken on any deviate data discovered in the samples and on any groups of N records concluded to have unacceptable quality. There may be three cases to consider. (1) Deviant data are found in the samples and the sampling plan finds the group of N records is acceptable. All deviant data found in the samples should be corrected. (2) Deviant data are found in the samples and the sampling plan finds the group of N records is unacceptable. This could mean checking and correcting the data in the N records. This could by checking each record, or by performing a second data entry and using software to compare the first and second files. (3) Deviant data may or may not be found in the samples, the sampling plan finds the group of N records is acceptable, but the results of the data checking suggests that improvements can be made to the all or some data in the group of N records.

Rules for Data Checking

UDS should establish rules to effectively and consistently apply the sampling plans. The random selections of records should be computer generated. Qualifications of the data checkers and the degree of participation of the data entry personnel should be established. UDS should define rules for checking the data against the paper records, for counting and recording the numbers and types of data deviations," and for determining when a data record is called errant. Data in some of the fields may have acceptable alternate values. If so, these should be listed and approved. The rules may be different for the three Sites.

UDS may want to be practical in establishing the data checking rules. As an example, consider possible rules for deciding when a record is counted as errant. Conservatively, a data record should be declared errant if any data deviation is found in the record. There may be cases where this is unnecessarily rigid. It may be more appropriate for minor, unambiguous spelling errors in –econdary information" fields to have less weight when deciding that a record is errant. To protect against pervasive minor carelessness, UDS might devise (i.e. invent) a minor deviation limit" with some value "k." Then, when the number of "minor deviations" equals k, that record is called errant. Counting "minors deviations" could start at zero for each new record, or it could be a cumulative total over different records. It could also be restricted to specific data fields. Note: This is not part of the statistical sampling plans. It is discussed to illustrate ways the plans could be adapted to better meet UDS' needs.

UDS may also choose to be practical in establishing the possible corrective actions. For example, check only specific data fields based on trends found in entry errors, rather than checking 100% of the data. This could save time, improve the quality of the data, and be used to prevent future entry errors.

Two Tables of Sample Sizes

Two tables of sample sizes are provided on pages 6 and 7. The sample sizes, n_1 and n_2, can be found when N is known. Table 1 gives samples sizes for acceptance numbers (c_1=0,c_2=4). Table 2 gives sample sizes for acceptance numbers (c_1=1,c_2=4). Either or both tables may be used. All of the sample sizes satisfy the two probability statements on page 1.

The Graph of Total Sample Size Versus N

The graph of the total sample size, $n=n_1+n_2$, on vertical axis, versus N, on horizontal axis, is moderately complex. The overall shape of the graph resembles the left half of the letter "U," cut from top to bottom, and rotated clockwise ninety degrees. The graph begins at the origin at the point (N=0, n=0). Sample sizes increase rapidly for small values of N, and the graph is nearly vertical. Increases in sample size slow for intermediate values of N, and the graph curves from nearly vertical to more horizontal. The sample size becomes a constant for some large value of N, where the graph is a true horizontal line.

The complexity of the graph lies in its sawtooth nature. Rather than being "smooth," the graph is a sawtooth, rotated, left half of a "U." Consider the graph n v. N for a single tooth. The sample sizes increase steadily as N increases. Then, with the an increase from N to N+1, the sample size drops radically. Then the pattern repeats. The size of the teeth are greatest for the smallest values of N, and decrease gradually as N increases. The teeth are gone when the graph becomes a true horizontal line. The sawtooth effect is a consequence of calculating probabilities using only integers values. The sample sizes, N, and the number of errant records in the data file are all integers.

Construction of the Sample Size Tables

The tables were constructed from the exact calculations of the sample sizes. Making short, easy to use tables required some simplifications. As a result, many of the tabled sample sizes are larger than the exact sample sizes.

1. The tables call for 100% data checking until the sample sizes stabilized enough to have reasonably consistent statistical properties over an interval of Ns.

2. The two samples sizes were set equal ($n_1 = n_2$) to reduce the chance of a user mistake. This increased some total sample sizes (n_1+n_2) by one for the (0,4) plan and by no more than two for the (1,4) plan.

3. The same sample sizes (n_1=n_2) were assigned to all of the Ns in an interval. This has eliminated the sawtooth effect from the tabled sample sizes. As a result, the tabled sample sizes exceed the exact sample sizes for many of the Ns in an interval. The largest increases occur in the intervals of the smaller Ns, where the sawtooth effect is greatest. For the (0,4) plan, n_1+n_2 is off by no more than two in intervals of the larger Ns, but can be off by as many as 10 in intervals of the smaller Ns. Similarly, for the (1,4) plan, n_1+n_2 is off by no more than four in intervals of the larger Ns, but can be off by as many as 14 in intervals of the smaller Ns.

Selection of the (c_1=0,c_2=4) and (c_1=1,c_2=4) Plans

The tables give sample sizes for the sampling plans having acceptance numbers (c_1=0,c_2=4) and (c_1=1,c_2=4). The 6/3/04 email attachment listed other combinations of acceptance numbers (please see page 8). Acceptance numbers (0,4) and (1,4) were selected for two reasons. First, they had the smallest sample sizes. And second, there could be concerns about using a sampling plan that can conclude N data records are acceptable when it is possible for "so many" errant records to be found in the samples. In fact, there may be concerns about the (0,4) and (1,4) plans. Any concerns can be addressed by referencing the "quality performance properties" given by the probability statements (1) and (2) on page 1, and by establishing appropriate data checking rules and the corrective actions.

References

The exact sample sizes were calculated using SAS[a]. The sampling probabilities are given by the statistical hypergeometric distribution. They were calculated using SAS' "PROBHYPER" function. These sample size calculations are presented in many newer books on statistical quality methods. "Double sampling plans for attributes" were developed for use during World War II when the calculations with the hypergeometric distribution were impractical. As a result, older quality control books give approximate methods to determine the sample sizes.

Duncan, Acheson J., *Quality Control and Industrial Statistics*, Richard D. Irwin, Inc., Homewood, IL, 1974.

Gibra, Isaac N., *Probability and Statistical Inference for Scientists and Engineers*, Prentice-Hall, Inc., Englewood Cliffs, NJ, 1973.

Schilling, Edward G., *Acceptance Sampling in Quality Control*, Marcell Dekker, Inc., New York, 1982.

---

[a] SAS is a registered trademark of The SAS Institute, Inc., Cary, NC. (Statistical Analysis System)

Operation of the Two Stage Sampling Plans

Definitions of the sampling plan symbols:

N = Number of data records to be checked for errors by the sampling plan.
n_1 = Number of records in the first sample.
n_2 = Number of records in the second sample, taken if needed.
c_1 = Lower acceptance number.
c_2 = Upper acceptance number.
e_1 = Number of errant records found in the first sample of n_1 records.
e_2 = Number of errant records found in the second sample of n_2 records, taken if taken.

To begin, determine N, the number of records to be checked for data entry errors by the sampling plan. Then, use N to look up the sample sizes, $n_1 = n_2$, using either Table 1 with (c_1=0 and c_2=4) on page 6, or Table 2 with (c_1=1 and c_2=4) on page 7.

Stage One Sampling:

Randomly select n_1 records from the N records. Use the USD rules to check the n_1 records for data errors. Count the number or errant records using the USD rules. This is the value of e_1.

Then,
if $e_1 \leq c_1$, conclude that the N records have acceptable quality,
or,
if $e_1 > c_2$, conclude that the N records have unacceptable quality,
or,
if $c_1 < e_1 \leq c_2$, move to stage two sampling.

Stage Two Sampling:

Randomly select n_2 records from the N-n_1 unchecked records. Use the USD rules to check the n_2 records for data errors. Count the number or errant records using the USD rules. This is the value of e_2.

Add the number of errant records found in the first and second samples. This is $e_1 + e_2$.

Then,
if $e_1 + e_2 \leq c_2$, conclude that the N records have acceptable quality,
or,
if $e_1 + e_2 > c_2$, conclude that the N records have unacceptable quality.

Corrective Actions:

Take appropriate actions on any errant data, on errant records found in the sample(s), on the N records that if they were concluded to have unacceptable quality according to the USD rules. And, if stated in the rules, take appropriate actions on any or all of the records that were concluded to have acceptable quality.

Table 1.
Sample Sizes for the Two Stage Sampling Plan
Having
Acceptance Numbers

$C_1 = 0$ and $C_2 = 4$

Number of Data Records: N	Sample Sizes: $n_1 = n_2$
216 or less	100% Checked
217 - 234	77
235 - 237	78
238 - 257	79
258 - 277	80
278 - 297	81
298 - 318	82
319 - 358	83
359 - 398	84
399 - 458	85
459 - 537	86
538 - 638	87
639 - 798	88
799 - 1098	89
1099 - 1718	90
1719 - 4178	91
4179 or more	92

Example: N=2580 data records will be checked for errors using the rules established by USD. Table 1 shows n_1=91 and n_2=91, with acceptance numbers c_1=0 and c_2=4.

Ninety-one records are randomly selected from the 2580 records and checked according to the set of rules. Two records are found to be errant. Using the notation on page 5, e_1=2. e_1 is compared to the acceptance numbers. Because 0<e_1 4, it is necessary to take the second sample.

Ninety-one additional records are randomly selected from the N-n_1=2489 unchecked records. One record is found to be errant, so e_2=1. Then, e_1+e_2=3.

Because e_1+e_2 4, the 2580 records are found to have acceptable quality. Appropriate actions are taken, as established by USD.

Table 2.
Sample Sizes for the Two Stage Sampling Plan
Having
Acceptance Numbers

$C_1 = 1$ and $C_2 = 4$

Number of Data Records: N	Sample Sizes: $n_1 = n_2$
217 or less	100% Checked
218 - 253	88
254 - 256	89
257 - 276	90
277 - 314	91
315 - 335	92
336 - 357	93
358 - 397	94
398 - 438	95
439 - 518	96
519 - 618	97
619 - 758	98
759 - 1018	99
1019 - 1538	100
1539 - 3178	101
3179 or more	102

Example: N=4576 data records will be checked for errors using the rules established by USD. Table 2 shows n_1=102 and n_2=102, with acceptance numbers c_1=1 and c_2=4.

One-hundred-two records are randomly selected from the 4576 records and checked according to the set of rules. One record is found to be errant, and e_1=1. Because e_1 c_1=1, the 4576 records are found to be acceptable.

Appropriate actions are taken, as established by USD.

For Reference:
Email Attachment J. Zolyniak to S. Barnes, 6/3/04.

Statistical Acceptance Sampling Plans
Values for Double Sampling Plans for Attributes

Double sampling plans for attributes are specified by the sample sizes for the two samples and two acceptance numbers: (n_1, n_2, c_1, c_2).

To review, suppose N data records are to be checked for entry errors. Randomly select n_1 data records and examine them for errors. If there are c_1 or fewer errors in the sample, the data file of N records is determined to have acceptable quality. If there are more than c_2 errors in the sample, the data file is determined to have unacceptable quality. If the number of errors in the first sample is more than c_1, but c_2 or less, a second sample is needed. Randomly select n_2 records from the unchecked N-n_1 records and examine them for errors. Then, if the total number of errors in the two samples is c_2 or less, the data file is determined to have acceptable quality; otherwise, the data file is determined to have unacceptable quality.

An Operating Characteristic (OC) curve@ is often used to describe performance properties of the sampling plan. This graph shows the probability that the sampling plan will find the N data records to have acceptable quality for different values of p, the true percentage of incorrect records in the file. We specified that the UDS acceptance sampling plan for checking data records should have the following two points on its OC curve. The Acceptable Quality Level (AQL)@ is 5% and this should have a probability of acceptance of at least 0.90. The Unelectable Quality Level (RQL)@ is 10% and this should have a probability of acceptance of no more than 0.10.

I was unable to determine values for (n_1, n_2, c_1, c_2) using the two points (5%, 0.90) and (10%, 0.10). The problem appeared to be with the AQL and its relation to the fixed point (0%, 1.0). An OC curve always has this point. When there are no incorrect records, the data file has 100% chance of being found acceptable. When I modified the AQL, I found values for (n_1, n_2, c_1, c_2). These sample sizes seemed rather small, in the order of 50-70 total. This may give us the opportunity to reduce the value of the RQL. Here are my findings. There was very little difference in the sample sizes for N=10,000 and N=20,000.

1. AQL = 0.5%, Probability of Acceptance >= 0.995
2. RQL = 5.0%, Probability of Acceptance <= 0.050

c_1 , c_2	n_1 , n_2	$n_1 + n_2$	c_1 , c_2	n_1 , n_2	$n_1 + n_2$
0 , 4	92 , 92	184	2 , 4	124 , 120	244
0 , 5	104 , 104	208	2 , 5	126 , 126	252
0 , 6	116 , 118	234	2 , 6	130 , 134	264
0 , 7	130 , 130	260	2 , 7	138 , 138	276
1 , 4	102 , 98	200[1]	3 , 4	152 , 136	288
1 , 5	110 , 110	220	3 , 5	152 , 184	336
1 , 6	120 , 120	240	3 , 6	154 , 132	286
1 , 7	130 , 132	262	3 , 7	156 , 146	302
3 , 8	160 , 154	314			

[1] Refinement of the computer programs resulted in a slightly higher sample size.