♻EPA

Data Quality Objectives Decision Error Feasibility Trials (DQO/DEFT)

User's Guide - Version 4.0
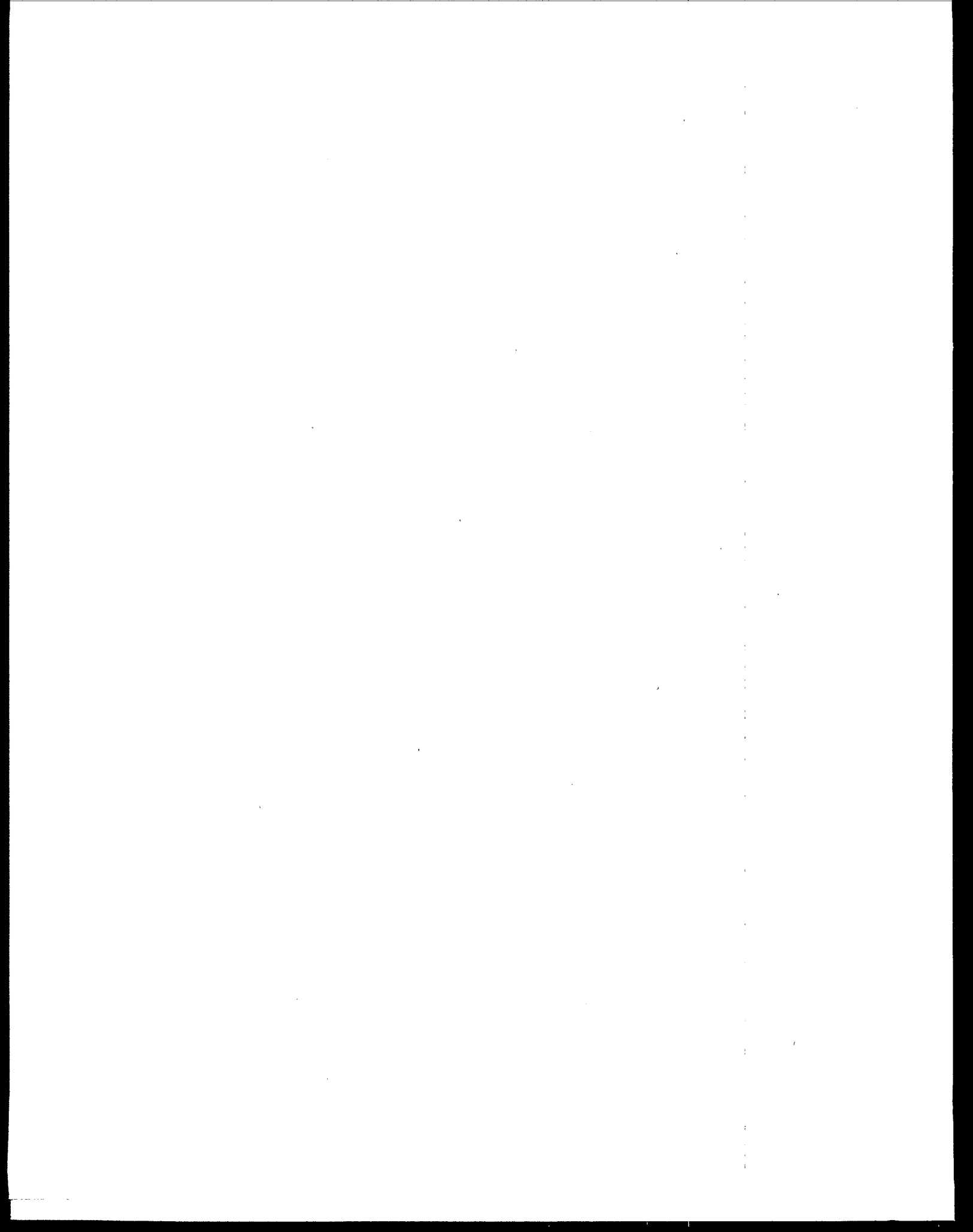
EPA QA/G-4D

♻EPA

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

USER'S GUIDE FOR DATA QUALITY OBJECTIVES DECISION ERROR FEASIBILITY TRIALS (DEFT) SOFTWARE

1. BACKGROUND

The two most intensive steps in the DQO Process are Step 6: Specify Tolerable Limits on Decision Errors and Step 7: Optimize the Design. During step 7, the entire set of DQO outputs is incorporated into a sampling design. If the DQO constraints are not feasible, it is necessary to iterate through one or more of the earlier steps of the DQO Process to identify a sampling design that will meet the budget and generate data that are adequate for the decision. This iteration can be time-consuming and costly. Therefore, the Decision Error Feasibility Trials (DEFT) software was developed to reduce the need for this iteration before implementing the final step of the DQO process. The DEFT software allows a decision maker or member of the DQO planning team to quickly generate cost information about several simple sampling designs based on the DQO constraints. Through this process, the planning team can evaluate whether these constraints are appropriate or feasible before the sampling and analysis design team begins developing a final sampling design in the last step of the DQO process (see Figure 1).

The DQO Process is described in the following two documents:

- *"Guidance for the Data Quality Objectives Process,"* EPA QA/G4. U.S. Environmental Protection Agency. 1994.

- *"The Data Quality Objectives Process for Superfund: Interim Final Guidance,"* EPA/540/R-93/071. U.S. Environmental Protection Agency. 1993.

The first document provides general guidance; the second provides guidance for Superfund applications. This software is designed to supplement these documents.

There is no easy method for developing a statistical sampling design. Factors such as environmental medium, parameter of interest, contaminant of interest, and sampling boundaries all affect the choice of a sampling design. For instance, volatile and non-volatile contaminants must be treated differently both in the field and in the laboratory. Sampling designs for soil, where samples can be randomly placed, are different from sampling designs for ground water, where sampling locations may be fixed. A composite sampling design is applicable for testing hypotheses concerning the mean; however, it is not applicable for testing hypotheses concerning percentiles. An optimal sampling design accounts for all these factors and others, yet is practical, feasible, and satisfies the DQO constraints. The DEFT software is not an expert system that will design an optimal (or even feasible) sampling design. It can be used only to evaluate the feasibility of the DQO constraints.

Decision makers (and the planning team) will be able to tailor the application of the DEFT software to their needs by entering basic information on the DQO constraints. The user will then be able to change DQO constraints such as limits on decision errors or the gray region, and evaluate how these changes affect the sample size for several basic sampling designs. The relations and values that are generated by the DEFT software can be used to set upper bounds on the sample size.
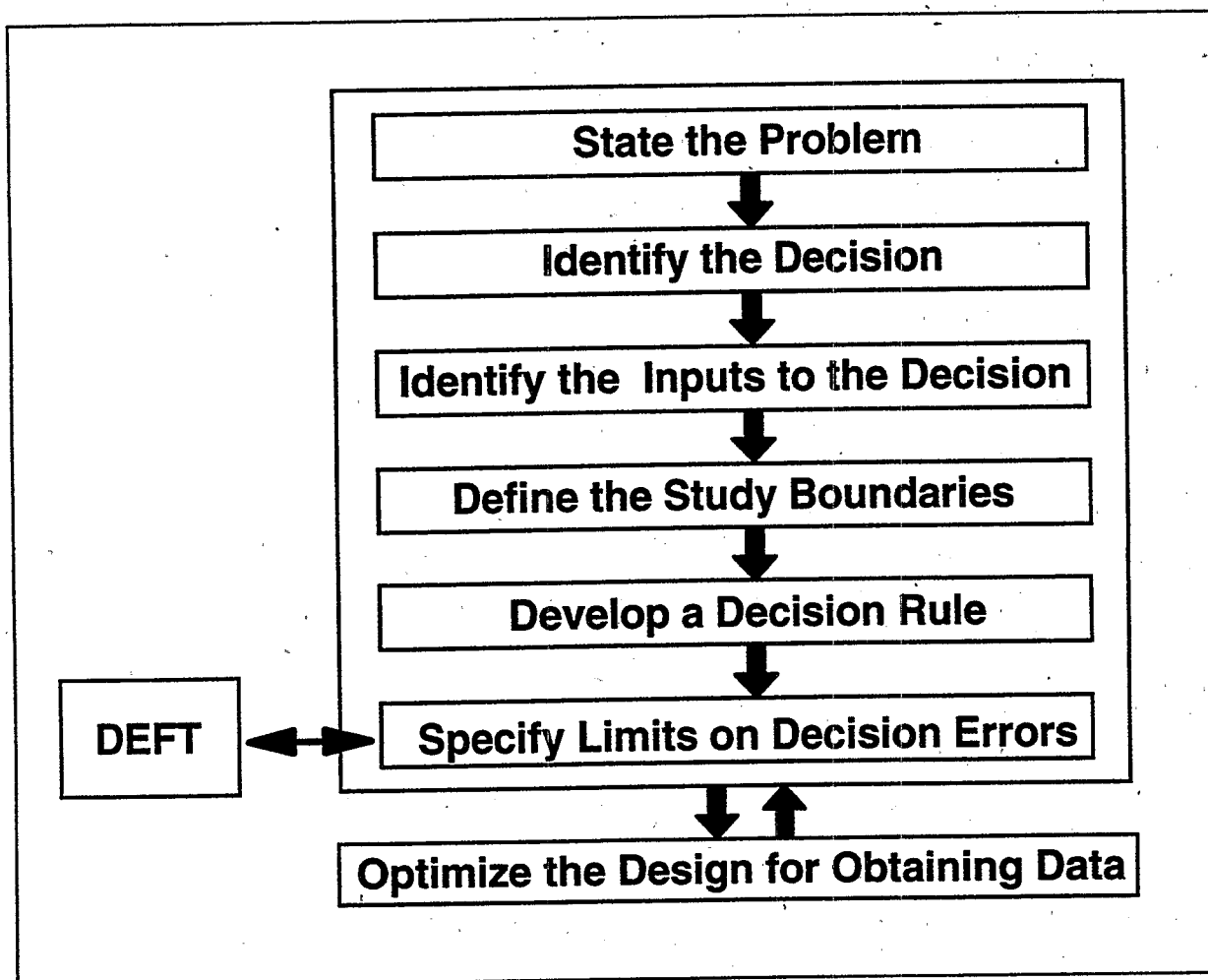
Figure 1. DEFT and the DQO Process

This user's guide contains detailed instructions on how to use the DEFT software and provides background information on the sampling designs that the software uses. However, the user's guide does not give in-depth instructions on the DQO Process or on the information the user is to input into the software. For more information on this information, consult the guidance listed above. The user's guide consists of five main chapters. Chapter 2 shows how to get started using the DEFT software and Chapter 3 tells how to use the DEFT software. Chapters 4-6 give detailed information on the DQO outputs that DEFT requires, the options available in DEFT, and the sampling designs implemented in DEFT.

2. INSTALLATION AND STARTUP

OVERVIEW OF VERSION 4.0

Version 4.0 of the DEFT software assumes that a population mean is being compared to a fixed action level (i.e., the action level is known). Future versions of DEFT will add additional statistical parameters and a variety of problems. The DEFT software also assumes that sample location can be randomized and there are no temporal issues. For example, in cases where drinking water samples are to be collected from wells whose locations are not randomly selected but instead selected based on hydrogeology, DEFT cannot be used.

INSTALLATION

The DEFT software can either be run from a floppy disk or the hard drive. Using the hard drive to run the software will speed up the start-up time and provide a directory for storing files saved using the software. To run the software using the hard drive, first install the software. To do this, insert the DEFT floppy disk into either drive 'a' or drive 'b'. Then type the following at the DOS prompt:

```
prompt>   a:<return>   (b:<return>)
prompt>   install a    (install b)
```

The installation program will install the DEFT software in the directory 'c:\deft'.

STARTING DEFT

If DEFT is installed on the hard drive of a computer, start the software by typing the following at the DOS prompt:

```
prompt>  c:<return>
prompt>  cd \deft<return>
prompt>  deft<return>
```

If DEFT is not installed on the hard drive of your computer, place the DEFT floppy disk into drive 'a' or drive 'b'. Then, at the DOS prompt type:

```
prompt>   a:<return>    (b:<return>)
prompt>   deft<return>
```

QUICK START: DEFT has a Quick Start option to allow the user to skip the entry screens and proceed directly to the Input Verification Screen to enter the DQOs. To implement this option, use either set of directions above and replace the final command with:

```
prompt>  deft q <return>
```

PRINTING THE USER'S GUIDE

A copy of this user's guide is contained with the software in the file "read.me". To view this user's guide, at the DOS prompt type:

prompt> type read.me |more <return>

To print another copy of the user's guide, the file "read.me" can be printed using any word processing package. Note that this copy of the user's guide does not contain any figures, tables, or mathematical equations.

FUTURE VERSIONS

Future versions of DEFT (Version 5.0+) will contain sampling designs for different parameters such as a percentile or proportion. Other planned changes to the DEFT software are mentioned in the relevant sections.

TROUBLESHOOTING

Sometimes a user may encounter an error message while running the DEFT software. Below are some errors that the user may encounter along with their solutions. If the DEFT software stops with an error message not listed below, contact the EPA Quality Assurance Management Staff (QAMS) at (202) 260-5763.

- "Error: can't set video mode." - If this error appears, use another monitor. The DEFT software selects the highest resolution available with the current hardware. The DEFT software should run on most EGA, CGA, and VGA color and monochrome monitors. However, it will not run on a MDPA (monochrome) monitor and on other adapters/monitors that do not support graphics modes.

- "Error: can't register fonts." - The file "helvb.fon" must be in either the DEFT directory (c:\deft) or on the floppy disk. If this file is missing, the DEFT software will not run. Contact QAMS for another copy of this file.

- "Error opening cover screen file." - The file "cover.scr" must be in either the DEFT directory (c:\deft) or on the floppy disk. If this file is missing, the DEFT software will not run. Contact QAMS for another copy of this file.

3. PROGRAM OVERVIEW

The DEFT software uses the outputs from DQO steps 1 - 6 and determines the feasibility of DQO constraints based on several simple sampling designs. This is done in three steps: 1) enter the information from the DQO outputs into DEFT, 2) verify and save the input information (the Input Verification Screen), and 3) summarize and analyze a sampling design in relation to the DQO constraints (the Design/DQO Summary Screen).

ENTRY SCREENS

The user is first prompted to enter information from the DQO outputs based on a series of entry screens. Information requested by the DEFT software includes:

- the parameter of interest,
- the minimum and maximum values (range) of the parameter of interest,
- the action level,
- the null and alternative hypotheses,
- the bounds of the gray region,
- an estimate of the standard deviation,
- sampling and analysis costs,
- probability limits on decision errors for the bounds of the gray region, and
- any additional limits on decision errors.

More information on these topics is contained in Chapter 4.

The DEFT software automatically starts with a simple random sampling design, so the information requested corresponds to this design. When requesting information, the DEFT software indicates in which DQO process step the information may be found. In addition, default values are offered for users who wish to use the software as a learning tool. Previous entries are summarized in the lower right-hand corner of the screen.

THE INPUT VERIFICATION SCREEN

Once the DQO constraints are entered, the DEFT software displays the Input Verification Screen (Figure 2). The Input Verification Screen is used to verify the inputs from the entry screens. Any incorrect values can be corrected at this time by pressing the highlighted letter corresponding to the topic. For example, press 'M' to change the minimum possible value for the parameter of interest. These keystrokes are summarized in Table 1 in Chapter 4. Once the information has been verified and corrected if necessary, the user may advance to the Design/DQO Summary Screen by entering 'Y' (or 'y').

The information on the Input Verification Screen is saved as the "Original DQOs", since this information represents the data quality objectives of the planning team. This gives the user the opportunity to select a sampling design, evaluate the performance of the design based on these Original DQOs, then change the DQOs to satisfy cost constraints. The user may then select a different sampling design and evaluate its performance based on the Original DQOs (i.e., the data quality objectives of the planning team), which may then also be adjusted to satisfy cost constraints.

```
        INPUT VERIFICATION SCREEN

   Are these values correct?  If so, press 'Y' to continue.
   Otherwise press the letter of the item to be changed.

   (M)inimum Concentration =  0.00
   Ma(X)imum Concentration = 100.00
   (A)ction Level = 50.00
   Gray Region = 50.00 - 75.00
   (S)tandard Deviation = 16.67
   (N)ull Hypothesis Ho: mean ≤ 50.00
   Cost of Analyzing Sample in a (L)aboratory = $1000.00
   Cost of Collecting a Sample in the (F)ield = $50.00


         Decision Error Limits:
      conc.              prob(error)    type
      ---        (1)        ---         F(+)
      ---        (2)        ---         F(+)
      50.00  (A) (3)       0.010        F(+)
      75.00  (B) (4)       0.010        F(-)
      ---        (5)        ---         F(-)
      ---        (6)        ---         F(-)
```

Figure 2. Example Input Verification Screen

For advanced users, future versions will allow the user to create a file containing DQO constraints and then start the DEFT software at the Input Verification Screen using the Quick Start option (see Chapter 2). Currently, the user may use the Quick Start option to start the DEFT software with the default values.

THE DESIGN/DQO SUMMARY SCREEN

After the user has left the Input Verification Screen, a Design/DQO Summary Screen is displayed (Figure 3). The summary screen contains information on the current sampling design and the DQOs, and includes several options for displaying, saving, and changing this information. The program always starts with a simple random sampling design. The user may then select different designs to explore. Possible sampling designs include:

1. Simple Random Sampling
2. Composite Sampling
3. Stratified Sampling

These sampling designs, along with information on the sample sizes and costs, are discussed in Chapter 6.

```
                    DESIGN/DQO SUMMARY SCREEN


    For the Sampling (D)esign of:  Simple Random Sampling

    Total Cost: $13650.00
          (L)aboratory Cost per Sample:  $1000.00
          (F)ield Cost per Sample:  $50.00

    (N)umber of Samples:  13


    Data Quality Objectives

    (A)ction Level:  50.00
    (S)tandard Deviation:  16.67
    Gray Region:  50.00 - 75.00
    Null Hypothesis: mean ≤ 50.00

              Decision Error Limits:
          conc.              prob(error)     type
          ---        (1)        ---          F(+)
          ---        (2)        ---          F(+)
          50.00   (A) (3)     0.010          F(+)
          75.00   (B) (4)     0.010          F(-)
          ---        (5)        ---          F(-)
          ---        (6)        ---          F(-)


    (G)raph      Sa(V)eFile      (O)riginal DQOs      E(X)it
```

Figure 3. Example Design/DQO Summary Screen

Cost and sample size information are provided on the current sampling design. The user may change the current sampling design or change the sample size at any time. The total cost also can be changed by adjusting the analytical or field sampling costs.[1] At this point, the DEFT software also verifies if the decision error limits are satisfied. If not, the limits that are not satisfied are marked "Not Satisfied". See Chapter 5 for more information on changing the sampling design, changing the sample size, and verifying the decision error limits.

The composite and stratified sampling designs require additional information from what was originally requested in DEFT for the simple random sampling design. For example, an estimate of the

---

[1]A future option will allow the user to constrain the total cost of the sampling design.

proportion of measurement variability to the total variability is required for the composite random sampling design. The first time a sampling design is selected, the user is prompted to enter the design-specific information. It is then possible to update this information in order to evaluate the DQO constraints in relation to this sampling design. Design-specific information is discussed in Chapter 6.

The Design/DQO Summary Screen shows information on the DQOs including the action level, the standard deviation, the decision error limits, and the analytic or field sampling costs. The user may change any of these values in the same manner they were changed on the Input Verification Screen. See Chapter 4 for more information on these variables. Note that it is not possible to change the minimum, maximum, and null hypothesis on the Design/DQO Summary Screen.

Several options are available on the Design/DQO Summary Screen for the user's convenience. These options include:

- Displaying the decision performance goal diagram with performance curve overlaid;
- Saving the current Design/DQO Summary Screen in a file; and
- Restoring the Original DQOs.

Detailed information on these topics is presented in Chapter 5.

To exit the DEFT software, the user may press 'X' (or 'x') on the Design/DQO Summary Screen.

4. DQO INFORMATION REQUIRED FOR DEFT

When running the DEFT software, the user is prompted to enter information from the first six steps of the DQO Process. This information is used to determine the appropriate sample size and associated cost of the sampling designs. This chapter of the user's guide describes the information that must be entered into DEFT. (Table 1 at the end of the chapter summarizes this information). The discussions below follow the order in which the user is prompted to enter information. The format of each discussion includes a description of the information to be entered, the step of the DQO process where this information is defined, possible default values, and the range in which the information must lie.

PARAMETER OF INTEREST

The parameter of interest is a descriptive measure of some characteristic or attribute of the statistical population. Possible parameters include:

1. the mean
2. a percentile
3. a proportion.

The current version of this program assumes that a population mean will be used; therefore, the user will not have to provide this information.

DQO STEP: The parameter of interest should have been identified in DQO Process Step 5: Develop a Decision Rule.

RANGE: Not applicable for the current version of the DEFT software.

DEFAULT: There will be no default. In future versions, the user must select a parameter of interest.

KEYSTROKE FOR INPUT VERIFICATION SCREEN: None

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: None

MINIMUM AND MAXIMUM VALUES (RANGE) OF THE PARAMETER OF INTEREST

If the parameter of interest is a population mean, estimates of the minimum and maximum possible values of this parameter are necessary for scaling and graphing purposes and for computing some default values. The range of the population mean should fall within the range of possible concentrations. These values are referred to throughout the rest of the DEFT software as the "minimum" and "maximum" concentrations.

DQO STEP: These values should have been identified in DQO Process Step 6: Specify Tolerable Limits on Decision Errors.

RANGE: There is no bound on the minimum concentration; however, the maximum concentration entered must be greater than the minimum concentration and less than 1,000,000,000.

DEFAULT: minimum concentration: 0.00
               maximum concentration: 100.0 + minimum concentration

KEYSTROKE FOR INPUT VERIFICATION SCREEN: Minimum - M
                                                        Maximum - X

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: None

ACTION LEVEL

The action level is a measurement threshold that provides the criterion for selecting among alternative actions. The current version assumes that the action level is fixed, such as a regulatory threshold or standard.

DQO STEP: The action level should have been identified in DQO Process Step 5: Develop a Decision Rule.

RANGE: The action level entered must be greater than the minimum concentration, but less than the maximum concentration.

DEFAULT: The default action level suggested is the midpoint between the minimum and maximum concentrations.

KEYSTROKE FOR INPUT VERIFICATION SCREEN: A

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: A

NULL AND ALTERNATIVE HYPOTHESES

The null (H_o) and alternative (H_a) hypotheses are used to identify which error is a false positive and which error is a false negative. This distinction is important to the calculation of the sample size. The current version of the DEFT software compares a mean to a fixed action level. Therefore, the user may select one of the following two choices to determine the null and alternative hypotheses:

1. H_o: mean $\geq$ Action Level   vs. H_a: mean $<$ Action Level

2. H_o: mean $\leq$ Action Level   vs. H_a: mean $>$ Action Level

Since the alternative hypothesis is simply the opposite of the null hypothesis, the DEFT software will only state the null hypothesis in future references.

DQO STEP: The null hypothesis should have been defined in DQO Process Step 6: Specify Tolerable Limits on Decision Errors.

RANGE: Valid entries are '1' or '2'.

DEFAULT: The default selection of the null hypothesis is selection '2'. (This default value was selected so that the decision performance goal diagram displays the expected power curve. Otherwise, the decision performance goal diagram displays the complement of the expected power curve.)

KEYSTROKE FOR INPUT VERIFICATION SCREEN: N

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: None

GRAY REGION

The gray region is a range of possible parameter values where the consequences of a false negative decision error are relatively minor. The gray region is bounded on one side by the action level and on the other side by that parameter value where the consequences of making a false negative decision error begin to become significant. Since the action level has already been entered, the DEFT software will prompt the user to enter the other bound of the gray region. The program will automatically determine whether this bound should be less than or greater than the action level.

DQO STEP: The gray region should have been defined in DQO Process Step 6: Specify Tolerable Limits on Decision Errors.

RANGE: If the null hypothesis is "H_o: mean $\geq$ Action Level," then the other bound of the gray region must be less than the action level and greater than the minimum concentration.

If the null hypothesis is "H_o: mean $\leq$ Action Level," then the other bound of the gray region must be greater than the action level and less than the maximum concentration.

DEFAULT: If the null hypothesis is "H_o: mean $\geq$ Action Level," the midpoint between the minimum concentration possible and the action level is the default value for the other bound of the gray region.

If the null hypothesis is "H_o: mean $\leq$ Action Level," the midpoint between the action level and the maximum concentration possible is the default value for the other bound of the gray region.

KEYSTROKE FOR INPUT VERIFICATION SCREEN: B

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: B

ESTIMATE OF STANDARD DEVIATION

An estimate of the standard deviation of the population of interest is necessary for computing sample sizes. The standard deviation is the square root of the variance. If there is no estimate available, then the default value will give a rough approximation of the standard deviation. Note that the default value is valid for the purposes of the DEFT software; i.e., determining the feasibility of the DQO constraints. The user should consult a statistician, however, before developing an estimate for use in the actual sampling design.

DQO STEP: An estimate of standard deviation may have been derived in DQO Process Step 3: Identify Inputs to the Decision.

RANGE: The standard deviation must be greater than zero and less than or equal to two times the range of the population parameter (i.e., the standard deviation must be less than or equal to two times the maximum concentration minus the minimum concentration).

DEFAULT: The default value used in DEFT is given by

$$(\text{Maximum Concentration} - \text{Minimum Concentration}) / 6$$

This default value gives a rough approximation of the standard deviation, but it should only be used if there is absolutely no other information available. This approximation is based on the range of the population, not the range of the population parameter.

If the planning team wishes to use this estimate, then an estimate of the range of the population should be entered in place of the range of the population parameter. (The population parameters considered in the DEFT software must all fall within the range of the overall population.) If the range of the population has not been entered, this may be corrected using the Input Verification Screen. Using the Input Verification Screen, first change the minimum and maximum concentrations, then change the standard deviation and use the new default value.

KEYSTROKE FOR INPUT VERIFICATION SCREEN: S

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: S

SAMPLING AND ANALYSIS COSTS

The average unit cost of analyzing a sample and the average unit cost of collecting a sample in the field are used to compute the total cost of a sampling design. The average cost of analyzing a sample is referred to as the "laboratory cost" and the average unit cost of collecting a sample is referred to as the "field cost" in the DEFT software.

DQO STEP: These costs may have been identified in DQO Process Step 3: Identify the Inputs to the Decision.

RANGE: Both the laboratory and field costs must be greater than or equal to 0.

DEFAULT: The default value of analyzing a laboratory sample is $1000. The default value for collecting a sample is $50. For the case where sample collection and measurement analysis are one process in the field, the user may enter the cost of this process as the laboratory cost and set the field cost equal to zero. Future versions will allow the user to specify a fixed set-up cost for field sampling, in addition to the existing unit costs per sample collected in the field.

KEYSTROKE FOR INPUT VERIFICATION SCREEN: Laboratory Cost - L
    Field Sampling Cost - S

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: Laboratory Cost - L
    Field Sampling Cost - S

PROBABILITY LIMITS ON DECISION ERRORS FOR THE BOUNDS OF THE GRAY REGION

Limits on the probability of committing a decision error at the bounds of the gray region must be specified. Therefore, the program will automatically prompt the user to enter the limits for these two points. These probability limits correspond to α, the probability of a false positive error, and β, the probability of a false negative error. The program will automatically determine which error is a false positive error, F(+), and which error is a false negative error, F(-), thus automatically determining which probability limit is α and which probability limit is β, based on the user's selection of the null hypothesis.

DQO STEP: The probability limits on decision errors for the gray region should have been defined in DQO Process Step 6: Specify Tolerable Limits on Decision Errors.

RANGE: All probabilities must be greater than 0 and less than 1. A probability greater than 0.5 (50%) of making a decision error, however, is similar to setting a probability limit of 0.5. Therefore, in the DEFT software, these probability limits must be greater than 0 and less than or equal to 0.5.

DEFAULT: The default value of the probabilities of making either decision error is 0.01.

KEYSTROKE FOR INPUT VERIFICATION SCREEN: 3 and 4

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: 3 and 4

ADDITIONAL LIMITS ON DECISION ERRORS

The DQO Process allows the planning team to set additional limits on decision errors besides those on the bounds of the gray region, although this is not necessary. Therefore, the DEFT software will allow the user to enter up to two additional limits below the lower bound of the gray region and up to two additional limits above the upper bound of the gray region.

This program will not use the additional limits to compute sample sizes. Therefore at times these additional limits may not be satisfied (which will be noted on the Decision Error Limits Table of the Design/DQO Summary Screen and the Decision Performance Goal Diagrams). If a probability limit is not satisfied, decrease the probability of error for the bound of the gray region or increase the width of the gray region.

To set additional limits on decision errors, the user is first prompted to set a limit below the lower bound of the gray region. If there are no additional limits below the gray region, simply hit the return key. If there is an additional limit, enter the concentration level where the probability limit will be specified. Once the concentration level has been entered, the user will be asked to enter the probability limit for this concentration. Tolerable probability limits should decrease as one moves away from the action level. This process will be repeated once again so that the user may enter a total of two additional decision error limits below the gray region.

After the user has entered either 0, 1, or 2 decision error limits below the gray region, the program will prompt the user to enter any additional decision error limits above the gray region. If there are no additional limits above the gray region, simply hit the return key. If there is an additional limit, enter the concentration level where the probability limit will be specified. Once the concentration level has been entered, the user will be asked to enter the probability limit for this concentration. Tolerable probability limits should decrease as one moves away from the action level. This process will be repeated once again, so that the user may enter a total of two additional decision error limits above the gray region.

DQO STEP: Additional decision error limits are identified in DQO Process Step 6: Specify Tolerable Limits on Decision Errors.

RANGE: The concentration level of any additional limits on decision errors below the gray region must be greater than the minimum concentration possible and less than the lower bound on the gray region.

The concentration level of any additional limits on decision errors above the gray region must be greater than the upper bound of the gray region and less than the maximum concentration possible.

All probabilities must be greater than 0 and less than or equal to 0.5.

DEFAULT: There are no default values. The user may or may not choose to enter additional limits.

KEYSTROKE FOR INPUT VERIFICATION SCREEN: Below the Gray Region - 1 and 2
                                             Above the Gray Region - 5 and 6

KEYSTROKE FOR DESIGN/DQO SUMMARY SCREEN: Below the Gray Region - 1 and 2
                                             Above the Gray Region - 5 and 6

Table 1: SUMMARY OF DQO INFORMATION USED IN DEFT

DEFT Input	DQO Process Step	Limits for Valid Entries	Default	Design/DQO Summary Screen	Input Verification Screen
Parameter of Interest	Step 5	None in Version 4.0	Mean.	None	None
Minimum Concentration (min) for Parameter of Interest	Step 6	None.	0.0	M	None
Maximum Concentration (max) for Parameter of Interest	Step 6	MAX > MIN	100 + MIN	X	None
Action Level (AL)	Step 5	MIN < AL < MAX	MIN + ½(MAX-MIN) (Midpoint between MAX and MIN)	A	A
Null and Alternative Hypothesis	Step 6	1. H_o: mean $\geq$ AL vs. H_a: mean < AL 2. H_o: mean $\leq$ AL vs. H_a: mean > AL	H_o: mean $\leq$ AL vs. H_a: mean > AL[a]	N	None
Gray Region (GR) (H_o: mean $\geq$ AL)	Step 6	MIN < GR < AL	MIN + ½(AL-MIN) (Midpoint between MIN and AL)	B	B
Gray Region (GR) (H_o: mean $\leq$ AL)	Step 6	AL < GR < MAX	AL + ½(MAX-AL) (Midpoint between AL and MAX)	B	B
Estimate of Standard Deviation (SD)	Step 3	0<SD<2*(MAX-MIN)	(MAX-MIN)/6[b]	S	S
Sampling Cost (SC)	Step 3	0 $\leq$ SC	$50	F	F
Analysis Cost (AC)	Step 3	0 $\leq$ AC	$1000	L	L
Probability of making false positive error at action level (α)	Step 6	0 < α $\leq$ 0.5	0.01	3, 4	3, 4
Probability of making false negative error at other end of gray region (β)	Step 6	0 < β $\leq$ 0.5	0.01	3, 4	3, 4

(continued)

DEFT Input	DQO Process Step	Limits for Valid Entries	Default	Design/DQO Summary Screen	Input Verification Screen
Additional Error Limits Above Gray Region. Concentration (CL) Probability (P)	Step 6	Limit of two additional entries. MIN < CL < GR or MIN < CL < AL $0 < p \le 0.5$	None.	1, 2	1, 2
Additional Error Limits Below Gray Region. Concentration (CL) Probability (P)	Step 6	Limit of two additional entries. GR < CL < MAX or AL < CL < MAX $0 < p \le 0.5$	None.	5, 6	5, 6

[a] This default value was selected so that the decision performance goal diagram displays the expected power curve. Otherwise, the decision performance goal diagram displays the complement of the expected power curve.

[b] This default value gives a rough approximation of the standard deviation; however, it should only be used if there is absolutely no other information available. This approximation is based on the range of the population, not the range of the population parameter. If the planning team wishes to use this estimate, an estimate of the range of the population should be entered instead of the range of the population parameter. (The population parameters considered in the DEFT software must all fall within the range of the overall population.)

5. OPTIONS AVAILABLE ON DESIGN/DQO SUMMARY SCREEN

On the Design/DQO Summary Screen, DEFT offers the user several options regarding the individual sampling designs. These options include changing the sampling design, constraining the number of samples, and verifying the decision error limits. Additional options include viewing the constraints and the design performance graphically, saving the Design/DQO Summary Screen to a file, and restoring the Original DQO constraints. A future option will allow the user to constrain the total cost of the sampling design.

CHANGING THE SAMPLING DESIGN (Keystroke 'S')

The following sampling designs are available in the DEFT software:

1. Simple Random Sampling
2. Composite Sampling
3. Stratified Sampling

The first time a sampling design is selected other than simple random sampling, additional information may be requested from the user. For more information on these sampling designs and the information required, see Chapter 6.

CONSTRAINING THE SAMPLE SIZE (Keystroke 'N')

Sometimes the user may know the total budget available for sampling and analysis. With this information, the user can determine the total number of samples that are affordable. If this is so, then the user may wish to determine what decision error limits are possible within this budget. To satisfy this situation, the DEFT software will allow the user to constrain the number of samples. The software will then adjust the probability of a false negative decision error (β) for the bound of the gray region. In most cases the sample size will then be equal to the value provided by the user. In some cases, however, the sample size may be slightly larger than the value provided by the user, due to the way DEFT performs the calculations under these conditions.

The user may enter any sample size greater than 1 and less than or equal to 1000. The stratified sampling design allows the user 1000 samples per stratum. (If the user has 4 strata, the total sample size must be less than or equal to 4000.) In some case, the sample size entered by the user may require the false negative decision error rate for the bound of the gray region to be greater than 0.50. In this case, the DEFT software will set the false negative decision error to 0.50 and determine the sample size necessary to meet this error rate. If this sample size is too large for the budget constraints, the user should expand the width of the gray region.

Note that after changing most of the DQO constraints (such as the action level or decision error limits), the program adjusts the sample size. This is still true even after the user has set the sample size. Therefore, if the user wishes to determine the limits on decision errors based on a fixed sample size over different assumptions of variability, the user must enter one estimate of variability, set the sample size, enter the next estimate of variability, then reset the sample size, and so on. The program will not automatically keep the entered sample size.
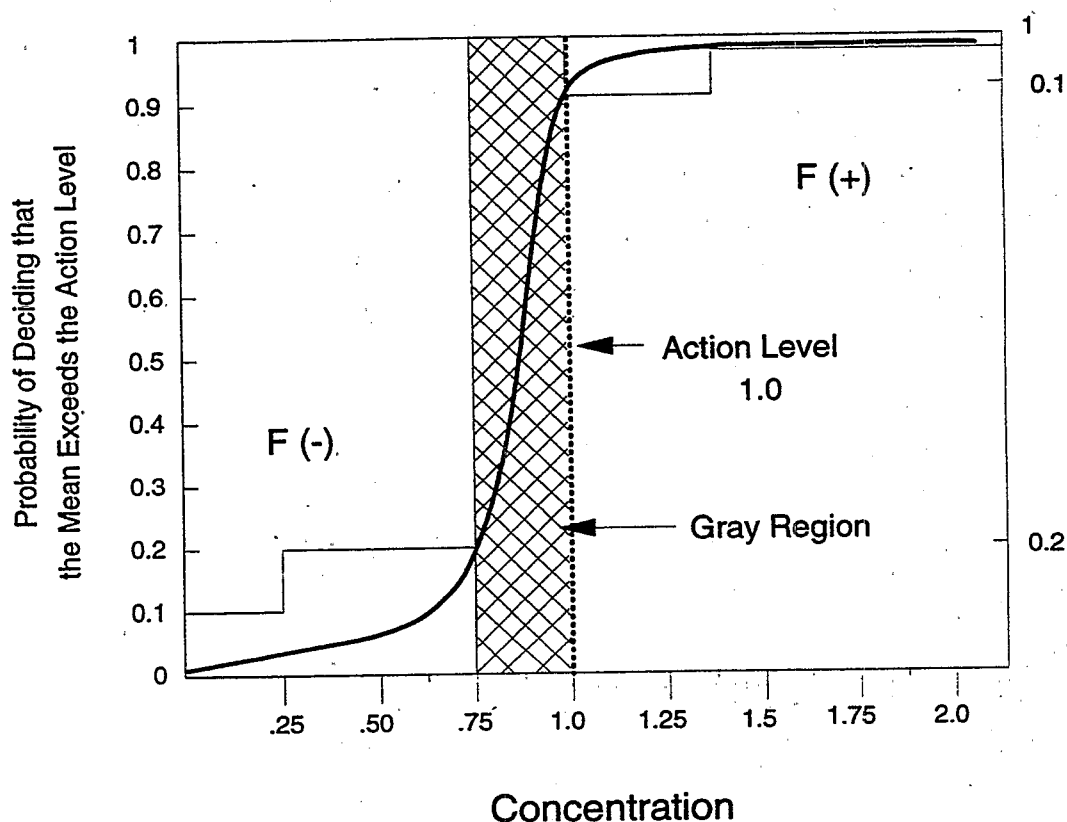
DECISION PERFORMANCE GOAL DIAGRAM

```
Simple Random Sampling
                            conc.      prob.       type
Action Level - 1.00         0.25       0.100       F(-)
                            0.75       0.200       F(-)
Cost - $13,650.00           1.00       0.100       F(+)
                            1.40       0.050       F(+)
Sample Size - 13


Press any key to return to the DESIGN/DQO Summary Screen.
```

Figure 4. Example Decision Performance Goal Diagram Screen with the Performance Curve (Using the Graph Option)

DISPLAYING THE DECISION PERFORMANCE GOAL DIAGRAM WITH THE PERFORMANCE CURVE (Keystroke 'G')

DEFT has an option available to view the DQOs and design performance graphically on a separate screen (Figure 4). This is done using a decision performance goal diagram with the performance curve overlaid. The performance goal diagram summarizes the gray region, the limits on decision errors, and the action level. In addition, information on the sample size and cost of the design are also summarized on this screen. An example of this screen is shown in Figure 3. After reviewing this screen, the user may press any key to return to the Design/DQO Summary Screen.

The performance curve is an approximation of the expected power curve or the complement of the expected power curve (depending on the null hypothesis). This curve can be used to determine how well a design performs in relation to the limits on decision errors. Note that the performance curve displayed by the DEFT software may grossly misrepresent the actual performance of the design on the false positive side of the gray region. In the current version of the DEFT software, a normal distribution is used to approximate the power curve, which is based on a non-central t-distribution. This is because the exact calculations are time-consuming without a math coprocessor installed in the computer. Future versions of this software will compute the exact power curve if a coprocessor is available.

The sample size reported by DEFT is always greater than or equal to 2 so that an estimate of the standard deviation can be calculated from the data collected. In this case, the performance curve may satisfy a more stringent false negative decision error rate at the bound of the gray region than that displayed by the software. If so, use the 'Constraining the Sample Size' option with a sample size of 2 to determine the exact decision error rate satisfied by the two samples.

SAVING THE CURRENT DESIGN INFORMATION AND THE DQOs IN A FILE (Keystroke 'F')

Once it has been determined that the DQO constraints are feasible for a sampling design, the user may wish to save the DQO constraints and design information to a file. Therefore, an option is available in the DEFT software which allows the user to save the current Design/DQO Summary Screen in an ASCII file. After selecting this option, the user will be prompted to enter a filename. DOS name conventions are used (up to 8 characters plus an optional 3-character extension separated by a period). If a file already exists under the name entered by the user, the program will ask the user to either select a new name or else to overwrite the existing file. Once a filename has been selected, the user may use this option to save any additional design and DQO information to this file. The additional information will be added to the end of the file.

RESTORING THE ORIGINAL DQOs (Keystroke 'O')

It is possible to restore the Original DQO constraints for simple random sampling specified on the Input Verification Screen and the DQO constraints entered when a composite design or a stratified design is first selected. This is useful for comparing variations of several sampling designs. For instance, the user will start with a set of DQO constraints and sampling design. The first sampling design may be too expensive to satisfy the DQOs, so the user may want to relax some constraints to

obtain a feasible sample size. After this is complete, the user may want to examine the performance of another sampling design using the Original DQO constraints. This option saves the user from re-entering the original information manually.

VERIFYING THE DECISION ERROR LIMITS

The sample size formulas used in the DEFT software guarantee that the decision error limits set on the bounds of the gray region are satisfied. The sample size formulas do not account for any additional decision error limits, however. Therefore, the DEFT software verifies that these additional limits are satisfied. If a limit is not satisfied, the limit is marked "Not Satisfied" in the Decision Error Limits Table.

Note that the performance curve may appear to show that a decision error limit is satisfied when it is not. This is because the performance curve, in the current version, only estimates the power curve. Note that the performance curve may grossly misrepresent the actual performance of the design on the false positive side of the action level. Therefore, the user should use the text indication in the Decision Error Limits Table to determine whether or not a limit is satisfied.

SAMPLE SIZE LIMITATIONS

The DEFT software has a sample size limitation of 1000 samples for a simple random sampling design and a composite random sampling design. There is a limit on the total sample size of 1000 times the number of strata for a stratified sampling design. For example, there is a limit of 4000 samples when the user has 4 strata. If the sample size required to meet the DQOs exeeeds the above limits, the DEFT software will set both the false positive error rate at the action level and the false negative error rate at the other bound of the gray region to 0.50. The user will then need to reset these error rates and make other adjustments to the DQOs (such as reduce the standard deviation or the increase the width of the gray region) in order to continue with the DQO constraint feasibility analysis.

6. SAMPLING DESIGN INFORMATION

SIMPLE RANDOM SAMPLING

The simplest probability sample is a simple random sample where every possible sampling point has an equal probability of being selected and each sample point is selected independently from all other sample points. Simple random sampling is appropriate when little or no information about a site is available. If some information is available, simple random sampling may not be the most cost-effective sampling design available.

The DEFT software automatically begins with the simple random sampling design.[2] Therefore, the information initially requested of the user is for a simple random sampling design. This information is described in Chapter 4 and in Table 1.

The DEFT software assumes that a t-test will be used to analyze the data. Therefore, the corresponding sample size formula is used in the computations:

$$n = \frac{\hat{\sigma}^2 (z_{1-\alpha} + z_{1-\beta})^2}{\Delta^2} + \frac{z_{1-\alpha}^2}{2} \qquad (1)$$

where: $\hat{\sigma}^2$ = estimated variance

z_p = the pth percentile of the standard normal distribution (from standard statistical tables)

Δ = the difference between the action level and the other bound of the gray region; and

n = the number of samples.

A derivation of this formula is contained in Appendix C of the *Guidance for the Data Quality Objectives Process* (EPA 1994). The sample size reported by the DEFT software is always greater than or equal to 2 so that an estimate of the standard deviation may be calculated from the data collected. Therefore, if the formula above yields a value less than 2, the DEFT software will automatically report a sample size of 2. In addition, if the sample size calculated is greater than 1000, the DEFT software will make adjustments to the false positive and false negative error rates. See the section on Sample Size Limits for more information (page 20).

The formula for computing the total cost of the simple random sampling design is:

$$Total\ Cost = n\ (\$\ per\ field\ sample\ +\ \$\ per\ laboratory\ sample) \qquad (2)$$

The performance curve calculations are also based on the suggested t-test. Currently, the software only approximates this performance curve instead of computing the exact curve because exact calculations are too time-consuming on most personal computers. As a result of this approximation, the performance curve may appear to show that a decision error limit is satisfied when it is not, especially on the false positive side of the gray region. Therefore, the DEFT software labels any

---

[2]The simple random sampling option may be used to develop upper bounds on the sample size for a randomized systematic sampling design (grid sampling with a random starting point).

decision error limit that is not satisfied as "Not Satisfied". This label should be used to determine whether or not a limit is satisfied instead of the performance curve. Future versions of DEFT will use the exact calculations instead of an approximation to the performance curve.

COMPOSITE SAMPLING

If analysis costs are high compared to sampling costs and the parameter of interest is a mean, then it may be appropriate to use composite samples to reduce the analysis costs. A composite sample is a physical mixing of two or more aliquots (grab samples) before analysis. The use of composite samples in association with a sampling design can be a cost-effective way to select a large number of sampling units and provides better coverage of the site without analyzing each unit.

The DEFT software uses composite samples with a simple random sampling design, referred to as "composite sampling". The software computes the number of composite samples, k, required to meet the DQOs based on a given number of aliquots, m, per composite sample. To determine the number of composite samples, an estimate of the ratio, r, of the relative standard deviation of measurement error to total standard deviation is required, along with the number of aliquots (m) to be contained within each composite sample. The user will be prompted to enter this information the first time a composite sampling design is selected. The user may then vary the number of aliquots within a composite sample by pressing 'C' and the ratio by pressing 'R'. The number of aliquots within a composite sample must be less than 100 and the ratio must be less than 1 and greater than 0. In addition, the user may also vary the total standard deviation as before. This information is summarized in Table 2.

The DEFT software assumes that a t-test will be used to analyze the data. The software then uses the corresponding sample size formula to determine the required number of composite samples (k) of size m to satisfy the current DQOs. The DEFT software assumes that the total variability across an exposure area can be represented as

$$\sigma^2_T = \sigma^2_X + \sigma^2_e, \tag{3}$$

where σ^2_T is the total variance across a population, σ^2_X is the variance inherent to the population (i.e., the true "natural" variability of contamination), and σ^2_e is the measurement error variance. Let

$$r = \hat{\sigma}_e / \hat{\sigma}_T. \tag{4}$$

If one forms k composite samples of size m, then the variance of the mean of the composites can be approximated as

$$\hat{V}(k, m) = [(1-r^2)/m]\hat{\sigma}^2_T + r^2\hat{\sigma}^2_T$$

$$= \hat{\sigma}^2_T [(1-r^2)/m + r^2]. \tag{5}$$

DEFT uses this estimate of $\hat{V}(k, m)$ in place of $\hat{\sigma}$ in equation (1) for simple random sampling. The resulting sample size is then the number of composite samples (k) of size m that should be selected in order to satisfy the DQOs. The sample size reported by the DEFT software is always greater than or equal to 2 so that an estimate of the standard deviation may be calculated from the data collected.

Therefore, if the formula above yields a value less than 2, the DEFT software will automatically report a sample size of 2. In addition, if the sample size calculated is greater than 1000, the DEFT software will make adjustments to the false positive and false negative error rates. See the section on Sample Size Limitations for more information (page 20).

The formula for computing the total cost of the composite sampling design is:

$$Total \ Cost = k \ [m \ (\$ \ per \ field \ sample) \ + \ (\$ \ per \ laboratory \ sample)] \quad \textbf{(6)}$$

The performance curve calculations are also based on the suggested t-test. Currently, the software only approximates this performance curve instead of computing the exact curve because exact calculations are too time-consuming on most personal-computers. As a result of this approximation, the performance curve may appear to show that a decision error limit is satisfied when it is not, especially on the false positive side of the gray region. Therefore, the DEFT software labels any decision error limit that is not satisfied as "Not Satisfied". This label should be used to determine whether or not a limit is satisfied instead of the performance curve. Future versions of DEFT will use the exact calculations instead of an approximation to the performance curve.

STRATIFIED SAMPLING

Stratified random sampling is used to improve the precision of a sampling design. To create a stratified sample, the study area is divided into two or more non-overlapping subsets (strata) that cover the entire site. Strata should be defined so that physical samples within a stratum are more similar to each other than to samples from other strata. Sampling depth, previous data and information about concentration level and previous cleanup attempts, and knowledge about contamination sources or activities can be used as the basis for creating strata. Once the strata have been defined, the DEFT software assumes each stratum will be sampled separately using a simple random design.

To estimate the sample size required for a stratified design, the DEFT software requires information regarding each individual stratum. There is a limit of four strata total in the DEFT software. For each stratum, the user will need to provide a weighing factor (weight) and an estimate of the standard deviation. The stratum weight is the proportion of the volume or area of the environmental medium contained in the stratum in relation to the total volume or area of the study site. The sum of the strata weights must be 1, so the program automatically computes the weight of the final stratum. The default weight corresponds to an equal weighing among the remaining strata. The estimated standard deviation for each stratum must be less than two times the range of the population parameter; the default value is the current estimate for the total standard deviation. This information is summarized in Table 2.

Due to space constraints on the Design/DQO Summary Screen, the strata information is not shown. To view and change the strata information, the user should select the "Display/Change Strata Information" option by pressing 'I'. The information on strata standard deviations and weights are then shown in a table. To change the strata standard deviations, press 'S' while viewing this table. To change the strata weights, press 'W' while viewing this table. To change the number of strata, press 'N' while viewing the table. If the number of strata is changed, the user will be prompted to enter estimates of the standard deviation and weights for each stratum.

The DEFT software assumes that a t-test will be used to analyze the data. Therefore, the corresponding sample size formula[3] (repeated for each stratum) is used in the computations:

$$n_h = [\sum_{h=1}^{L} W_h \hat\sigma_h] \ [\frac{(z_{1-\alpha} + z_{1-\beta})^2}{\Delta^2}] \ W_h \hat\sigma_h \qquad (7)$$

where:

n_h = the number of samples for stratum h;
L = total number of strata;
W_h = weight for stratum h;
$\hat\sigma_h$ = estimated standard deviation for stratum h;
Δ = the difference between the action level and the other bound of the gray region; and
z_p = the pth percentile of the standard normal distribution (from standard statistical tables).

The sample size reported by the DEFT software is always greater than or equal to 2 so that an estimate of the standard deviation may be calculated from the data collected in each stratum. Therefore, if the formula above yields a value less than 2, the DEFT software will automatically report a sample size of 2. This means that the minimum sample size for a stratified design is equal to two times the number of stratum. If the sample size calculated is greater than 1000 times the number of strata, the DEFT software will make adjustments to the false positive and false negative error rates. See the section on Sample Size Limitations for more information (page 20). In addition, this sample size formula assumes that the costs of sampling each stratum are the same. If not, see Chapter 6 of *Methods for Evaluating the Attainment of Cleanup Standards* (EPA 1989) for a sample size formula corresponding to unequal costs.

The formula for computing the total cost of the stratified sampling design is:

$$Total\ Cost = \sum_{h=1}^{L} n_h(\$ \ per\ field\ sample + \$ \ per\ laboratory\ sample) \qquad (8)$$

The performance curve calculations are also based on the suggested t-test. Currently, the software only approximates this performance curve instead of computing the exact curve because exact calculations are too time-consuming on most personal computers. As a result of this approximation, the performance curve may appear to show that a decision error limit is satisfied when it is not, especially on the false positive side of the gray region. Therefore, the DEFT software labels any decision error limit that is not satisfied as "Not Satisfied". This label should be used to determine whether or not a limit is satisfied instead of the performance curve. Future versions of DEFT will use the exact calculations instead of an approximation to the performance curve.

---

[3]This sample size formula assumes that the standard deviation is known. Therefore, when the standard deviation is estimated and the calculated sample size is small, consider increasing the sample size by 2 or 3 samples per stratum.

Table 2: SUMMARY OF DESIGN INFORMATION USED IN DEFT

Sampling Design	Design Information	Limits	Default	Input Verification Screen
Simple Random Sampling	Number of samples (n)	$2 \le n \le 1000$	None	N
Composite Sampling	Number of composite samples (k)	$2 \le n \le 1000$	None	N
	Number of aliquots in composite (m)	$1 < m < 100$	4	C
	Ratio (r) of measurement SD to total SD	$0 < r < 1$	0.10	R
Stratified Sampling	Number of strata (L)	$2 \le L \le 4$	2	I L
	Total number of samples (n)	$2 \le n \le 1000L$	None	N
	Stratum weights (W_h)	$0 < W_h < 1$	Equal weight among strata.	I W
	Stratum standard deviation (σ_h)	$0 < \sigma_h < 2(max-min)$	σ_{Total}	I S

REFERENCES

U.S. Environmental Protection Agency. 1994. *Guidance for the Data Quality Objectives Process.* EPA QA/G4, Office of Research and Development.

U.S. Environmental Protection Agency. 1993. *The Data Quality Objectives Process for Superfund: Interim Final Guidance.* EPA/540/R-93/071, Office of Emergency and Remedial Response.

U.S. Environmental Protection Agency. 1989. *Methods for Evaluating the Attainment of Cleanup Standards: Volume I: Soils and Solid Media.* EPA/230/02-89-042, Office of Policy Planning and Evaluation.