EPA-450/4-89-015

Guidance On Applying The Data Quality Objectives Process For Ambient Air Monitoring Around Superfund Sites (Stages I & II)

By

Research Triangle Institute
Research Triangle Park, NC 27709

EPA Contract No. 68-02-4550

EPA Project Officer: Darryl von Lehmden
Atmospheric Research And Exposure Assessment Laboratory

Office Of Air Quality Planning And Standards
Office Of Air And Radiation
U. S. Environmental Protection Agency
Research Triangle Park, NC 27711

August 1989

ACKNOWLEDGEMENTS

We would like to thank Stanley Sleva of the Office of Air Quality Planning and Standards, as well as Work Assignment Officer Jane Leonard, for their efforts and guidance on this project. In addition, the following individuals provided helpful comments and suggestions:

Thomas Pritchett, OSWER
Peter Kahn, Region I
Jody Hudson, Region VII
Thomas Curran, OAQPS
Joseph Pagdett, OAQPS

PREFACE

This document, Guidance for Applying the Data Quality Objectives Process for Ambient Air Monitoring Around Superfund Sites, provides direction and assistance on how the data quality objectives (DQO) process is used to design an ambient air monitoring system around a Superfund site that will be adequate for the intended use of the data. It is intended to serve as a bridge between the Quality Assurance Management Staff (QAMS) DQO guidance and actual application of the DQO process at Superfund sites. Specifically, this document was written to aid the Remedial Project Managers, Enforcement Project Managers, and the EPA Regional and Superfund contractor personnel responsible for ambient air sampling and analysis at Superfund sites to carry out their jobs in an efficient and effective manner.

DQOs are statements of the level of uncertainty that a decision maker is willing to accept in results derived from environmental data. The DQO process consists of three stages. This document gives an example of Stage I (preliminary definition of the decision) and Stage II (refinement of the decision and requirements) for monitoring ambient air quality during remedial action at a hypothetical Superfund site. Stage III of the DQO process, to be completed at a later date for this hypothetical site, involves the design of an ambient air monitoring system capable of achieving the DQOs.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

SECTION 1.0
INTRODUCTION

1.1 PURPOSE

The purpose of this document is to provide direction for and assistance in using the data quality objectives (DQO) process to design ambient air monitoring systems around Superfund sites that will generate data of adequate quality for use in decision making at the specified levels of certainty. This guidance is also meant to assist the reader in adapting the DQO process, as presented by the Quality Assurance Management Staff (QAMS) in Development of Data Quality Objectives, Description of Stages I and II, to site-specific problems: it gives an example of Stages I and II of the DQO process for monitoring ambient air quality during remedial action at a hypothetical site.

1.2 OVERVIEW OF THE DATA QUALITY OBJECTIVES PROCESS

DQOs are statements of the level of uncertainty that a decision maker is willing to accept in decisions based upon environmental data. Developing DQOs should be one of the first steps in initiating an environmental data collection program to be conducted by or for the U.S. Environmental Protection Agency (EPA). The DQO process helps decision makers, data users, and data generators communicate clearly with each other about the purposes for which environmental data will be used, the resources which can be made available for the effort, and the level of quality required of the results to be derived from the environmental data. For planning an appropriate ambient air monitoring program at Superfund sites, communication between decision makers and data users, such as Remedial Project Managers (RPMs) and Enforcement Project Managers (EPMs), and data generators, such as personnel responsible for ambient air sampling and analysis or State personnel, is important. Equally important is the interactive role of the decision maker and the technical staff in defining the overall goal(s) (or decisions) of a given monitoring program. Not only can a project manager initially select a goal that may not be technically achieved (e.g., 10^{-6} risk for bis-2-chloromethyl ether with 95% confidence), he may need significant input from his technical staff just to properly define his initial decision. For example, most project managers would not be aware of any differences in the quality of risk determinations based upon single event, "worst case" concentrations versus risk determinations based upon modeled, total yearly exposures using multiple monitoring events, and historical meteorological data.

The DQO process consists of three stages with several steps. Table 1 summarizes the DQO process in terms of purpose and the individual(s) serving in the lead role for each stage. The decision maker takes the lead role in defining the decision in Stage I, and the program and technical staff lead in establishing qualitative and quantitative constraints in Stage II. The process described in the first two stages results in proposed DQOs with accompanying specifications (constraints). The specifics of monitoring or modeling system design are not explicitly considered until Stage III. In this stage, alternative designs for collecting the required data are developed and quantitatively evaluated, leading to the selection of the monitoring design which ensures that the DQOs will be met in an economical manner. The DQO process is meant to be iterative among all stages. If the DQOs cannot be satisfied with the given resources and other constraints, the decision maker must decide whether to allocate more resources, relax the DQOs, or change some other aspect of the monitoring design.

DQOs for a monitoring program are achieved by executing an experimental design which is appropriate for the expected measurement data quality. DQOs differ from measurement quality objectives (sometimes referred to in quality assurance project plans as data quality indicator goals) in that they are limits for the OVERALL uncertainty of a PROJECT's results. Measurement quality objectives, such as precision and accuracy, are limits for the uncertainty of specific MEASUREMENTS. A DQO is expressed as the probability of making a wrong decision; a measurement quality objective is expressed as a desired value of precision, accuracy, completeness, or representativeness. Although they describe quality at different levels, DQOs and measurement quality objectives are directly and quantitatively related. During Stage III of the DQO process, monitoring program designs are evaluated for their ability to satisfy DQOs, given the expected precision, accuracy, and completeness of the measurement data. For example, if the precision of one sampling and analysis procedure is not sufficient to meet a DQO, then the procedure must be changed to require more measurements to improve its precision or a more precise procedure must be identified and used. During or following execution

TABLE 1. DQO PROCESS SUMMARY

STAGE	I		II	III
PURPOSE	Define Decision and Time Frame of Decision	Establish Feasibility of the Decision	Establish Qualitative and Quantitative Constraints	Design Data Collection Program To Meet Constraints
LEAD ROLE	Decision Maker	Technical and Program Staff	Technical and Program Staff	Technical Staff

of the experimental design, the monitoring program is again evaluated for its satisfaction of DQOs by using actual precision, accuracy, and completeness data obtained from the quality assurance/quality control (QA/QC) program. Thus, DQOs and measurement quality objectives are reconciled in both the planning of the experimental design to ensure that DQOs will be met and by analysis of QA/QC data from the monitoring program to assure the decision maker that the DQOs were met.

1.3 CONSIDERATIONS FOR AMBIENT AIR MONITORING AROUND SUPERFUND SITES

The physical characteristic of the site itself is one consideration in designing an efficient and effective air quality monitoring program. For example, whether or not the site acts as a point source or area source, and whether or not the air emissions are at ground level or from some elevated point influences the resultant monitoring strategy. Another consideration is seasonal change. If volatile organic compounds are of concern, hot summer months may increase the potential for emissions and thus require a more intense and quantitative monitoring effort. On the other hand, cold winter months, especially with snow and ice covering the ground, may minimize or preclude any emissions and thus reduce the monitoring requirements. Another important consideration is the time of day during which air samples are taken. For example, diurnal variations in downwind concentrations due to changes in the meteorological conditions such as temperature and solar radiation changes, are rarely considered even though they may be the overriding consideration. Too many times EPA investigates odor complaints by conducting air monitoring studies during daylight hours despite the vast majority of complaints being logged in the early evenings or early mornings.

A general air quality monitoring strategy for a Superfund site is to first characterize the site qualitatively and semi-quantitatively using a quick screening procedure. Then, if results from the screening procedure indicate a potential problem either in terms of the types of pollutants present or their concentration levels, a more rigorous procedure is used to quantitatively characterize and monitor the site.

The routine monitoring program during the remedial response activity may have several levels of intensity. For example, if the initial screening procedure does not indicate a problem, then scheduled periodic screenings may be all that is needed. If, at the other extreme, the initial or any subsequent screening indicates that pollutant concentrations are near the level of concern, then a more rigorous monitoring effort is required to increase the probability of detecting an exceedance and of being able to respond to that exceedance in a timely manner. The more rigorous procedure could involve the on-site use of compound-specific measurements in near real-time using portable gas chromatographs, for example. More specific considerations regarding instrumentation and exposure limits are discussed in the following paragraphs.

Paramount to the selection of instrumentation for ambient air monitoring is a knowledge of the following:

- pollutants of interest and their expected concentration ranges at points where they will be measured;
- required data quality (screening, qualitative, or quantitative) to meet the DQOs;
- available sampling and analysis procedures and their turnaround times (real-time results may be necessary in some situations, while a delay of 48 hours or longer to receive laboratory results may be acceptable in other cases); and
- available resources including funds, personnel (training, availability, and level of experience), and equipment.

All of the above items are addressed through application of the DQO process.

The first consideration for the selection of specific instrumentation for the monitoring system is the pollutants that may be emitted from the site. The list of compounds of interest may include all that could create a public health problem at the site or only indicator chemicals selected on the basis of quantities present at the site, level of hazard, volatility, persistence, and severe odor. The concentration levels of concern may be based on regulations such as applicable or relevant and appropriate requirements (ARARs), or health risk-based calculations, as given in the Superfund Public Health Evaluation Manual. ARARs may be Federal, State, or local regulations.

Volatile organic compounds (VOCs) are one class of compounds of special interest at Superfund sites and are used in the example in this document. There are no Federal ambient air regulations for VOCs. If a site with a potential ambient air VOC problem is not covered by State or local ARARs, health risk-based calculations can be used if appropriate toxicity constants are available. During periods of remedial activity, when short-term exposures are the concern, the health risk-based calculations are not directly applicable because only long-term toxicity constants are available. Thus, in these instances, concentration levels of concern may need to be based on exposure standards such as those promulgated under the Occupational Safety and Health Administration (OSHA), i.e., TLV/10, American Conference of Governmental Industrial Hygienist (ACGIH), or an extrapolation to short-term exposures from the health risk-based calculation.

The available sampling and analysis schemes that may be applicable for ambient air monitoring for VOCs around Superfund sites include the following:

- Colorimetric detection tubes provide quick turnaround time (i.e., the time between sample collection and receipt of analytical results) and are inexpensive. Using these tubes for screening may be appropriate if the concentrations are far above the levels of concern to provide for a quick decision to halt remedial activities and/or evacuate the affected population. If this method indicates a concentration near the level of concern, a more rigorous sampling and analysis scheme may be needed since this method is not especially accurate. Many tubes are cross sensitive to other compounds and may give false positives. These measurements made on-site can be used with appropriate meteorological data in models to predict receptor concentrations.

- Portable instruments (those that can be carried by an individual) such as flame ionization detectors and photoionization detectors (e.g., HNU meters) may be used if quick turnaround time is needed and if the levels of concern are in terms of total organic concentrations instead of compound-specific. If these instruments indicate a potential problem, more compound-specific analyses will usually be needed. The accuracy of these instruments for total organics depends somewhat on the frequency of calibration. These instruments can be used for measurements at the source, the fenceline, and/or the receptors.

- More sophisticated transportable on-site instruments (those that can be moved in a van, but not carried by an individual) such as gas chromatographs (GCs) are compound-specific and provide almost real-time results. These systems should be used when the concentration levels are near the levels of concern, and good precision and accuracy are needed to minimize the risk of making a wrong decision. If properly used and calibrated, they provide data quality adequate for most decision making. These systems can be moved to follow a plume and evaluate hot spots at the fenceline or near the receptors. They require highly trained staff to operate and maintain.

- Off-site laboratory instrumentation, including GC, gas chromatography/mass spectrometry (GC/MS), and high performance liquid chromatography (HPLC), can be used to analyze samples collected on-site and transported to the laboratory. This instrumentation usually will provide the most reliable compound-specific data, but does not provide real-time or almost real-time data. Decisions to implement management practices to control ambient air emissions, to halt remedial activities, to evacuate the receptor population, or to continue with the remedial activities may be made based on this type of data, depending on the concentration level and the data turnaround time. These types of laboratory analyses may be used to confirm results from other methods discussed above.

Modeling may also be used to make decisions at Superfund sites, as when source measurements and meteorological data are used in a model to predict receptor concentration. Final selection of which sampling, analysis, and modeling methods to use should only be made after considering all of the issues raised in Stages I and II of the DQO process.

1.4 USE AND ORGANIZATION OF THIS DOCUMENT

This document should aid the Remedial Project Managers, Enforcement Project Managers, and the EPA Regional and Superfund contractor personnel responsible for ambient air sampling and analysis at Superfund sites. Specifically, the intent of this document is to serve as a bridge between DQO guidance documents and the site-specific DQOs that are required for ambient air monitoring at Superfund sites.

The reader should be familiar with the following:

- the QAMS guidance: Development of Data Quality Objectives, Description of Stages I and II
- the two volumes by the Office of Solid Waste and Emergency Response: Data Quality Objectives for Remedial Response Activities
- the two DQO papers prepared by OAQPS: Data Quality Objectives for the Toxic Air Monitoring Systems and Data Quality Objectives for the Urban Air Toxic Monitoring Program
- the four volumes of Procedures for Conducting Air Pathway Analyses for Superfund Applications

The remainder of the document is organized as follows. Sections 2.0 and 3.0 present Stages I and II, respectively, of the DQO process for the hypothetical site. The last subsection in Section 3.0 is a summary of the DQOs for this site. Appendix A is a statistical discussion on false positive and false negative errors in terms of the DQO process. Appendix B is a glossary of Superfund program acronyms, relevant statistical terms, and terms from the DQO process.

SECTION 2.0
STAGE I: PRELIMINARY DEFINITION OF THE DECISION

During Stage I of the DQO process, the decision maker provides a preliminary definition of the decision. The definition includes a description of the problem which in this case is a Superfund site, identification of the decision makers and key data users and their roles in this program, the decision(s) and associated actions, information and measurement data requirements and uses, consequences of making an incorrect decision, a ranking of the seriousness of the errors, and the resources available for the program.

2.1 PROBLEM DESCRIPTION

This hypothetical site, Superdump, is approximately 180 acres and is located on the outskirts of an urban area. Figure 1 is a map of the site. Approximately 20,000 drums containing wastes (mainly solvents) from local industries were buried on the site from 1959 until 1980. Some of the drums have leaked, and contamination of soil and groundwater is possible. Leakage or accidental spills may also create an ambient air problem from vaporization, especially during the summer months when daytime temperatures are frequently above 90°F. The drums contain mainly spent solvents and have RCRA waste codes of F001, F002, F003, F004, F005, and F006, according to site records. The possibility that additional, undocumented compounds may be present at the site must also be considered. Most of the drums were buried at a depth of about 2 meters and in a single layer. The 3-acre area marked "A" in Figure 1 designates where the drums were buried. The area marked "B" has a State-operated air monitoring station which has been operating for a number of years. Its location adjacent to Superdump is coincidental, but historical meteorological data such as wind direction, wind speed, rainfall, and temperature will be useful in designing the monitoring network. Any chemical measurements made at the station will be of limited usefulness in detecting emissions from the site because the station is not upwind or downwind of the site. Table 2 provides relevant site information.

A Remedial Investigation/Feasibility Study (RI/FS) for Superdump and plans for remediation have been completed. The remedial action selected following the RI/FS is to dig up the buried drums and, after stabilization

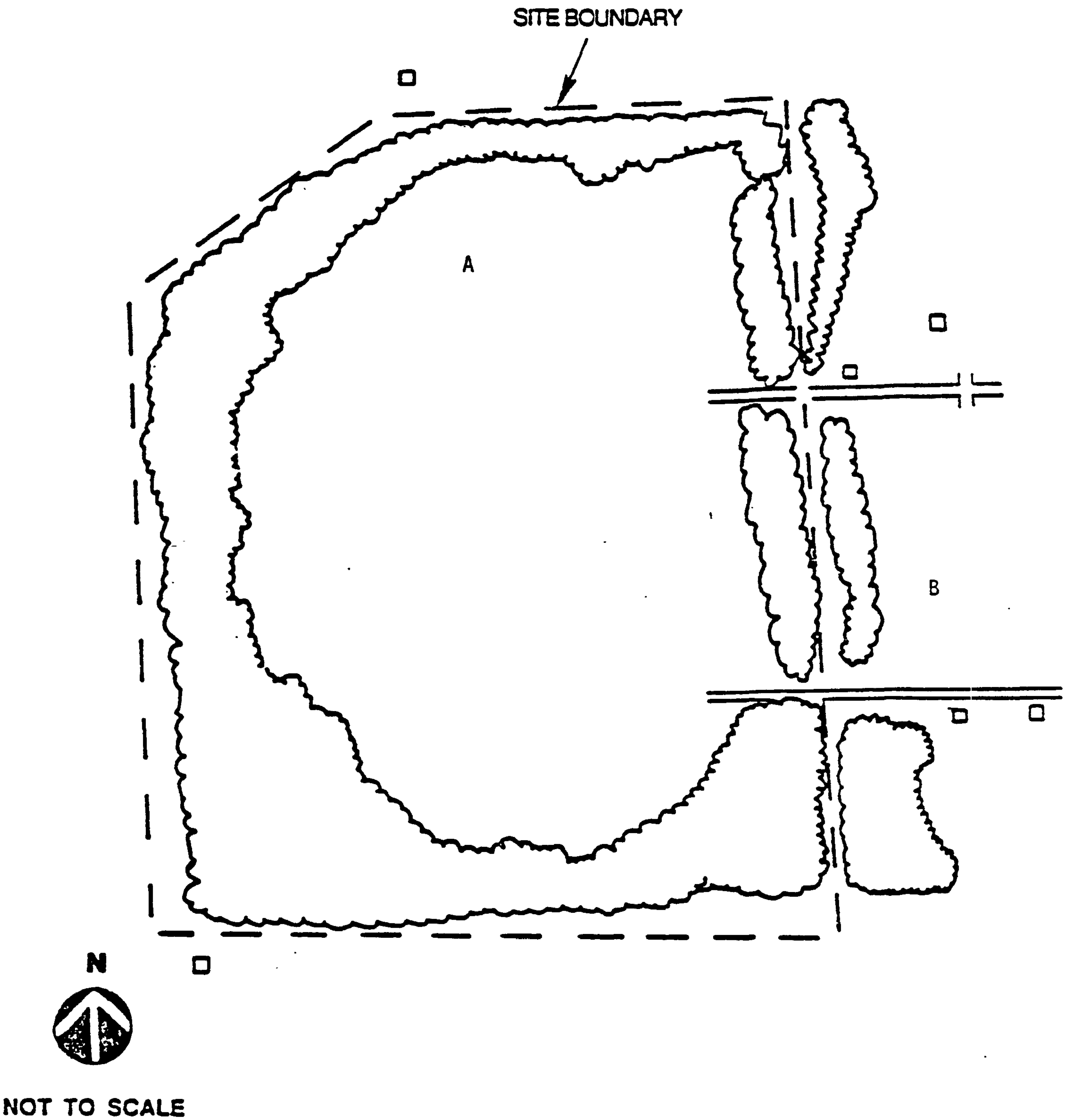


Figure 1. Superdump site map.

TABLE 2. SUPERDUMP SITE INFORMATION

Site size	180 acres
Contaminated area	3 acres
VOC concentration in soil	0 to 500 parts per million
VOC concentration in groundwater	0 to 30 parts per million
Site terrain	Level.
Nearest downwind receptor	100 meters
Average wind speed	2 meters/second
Prevailing wind direction	From the northeast quadrant in winter and from the southwest quadrant in summer.

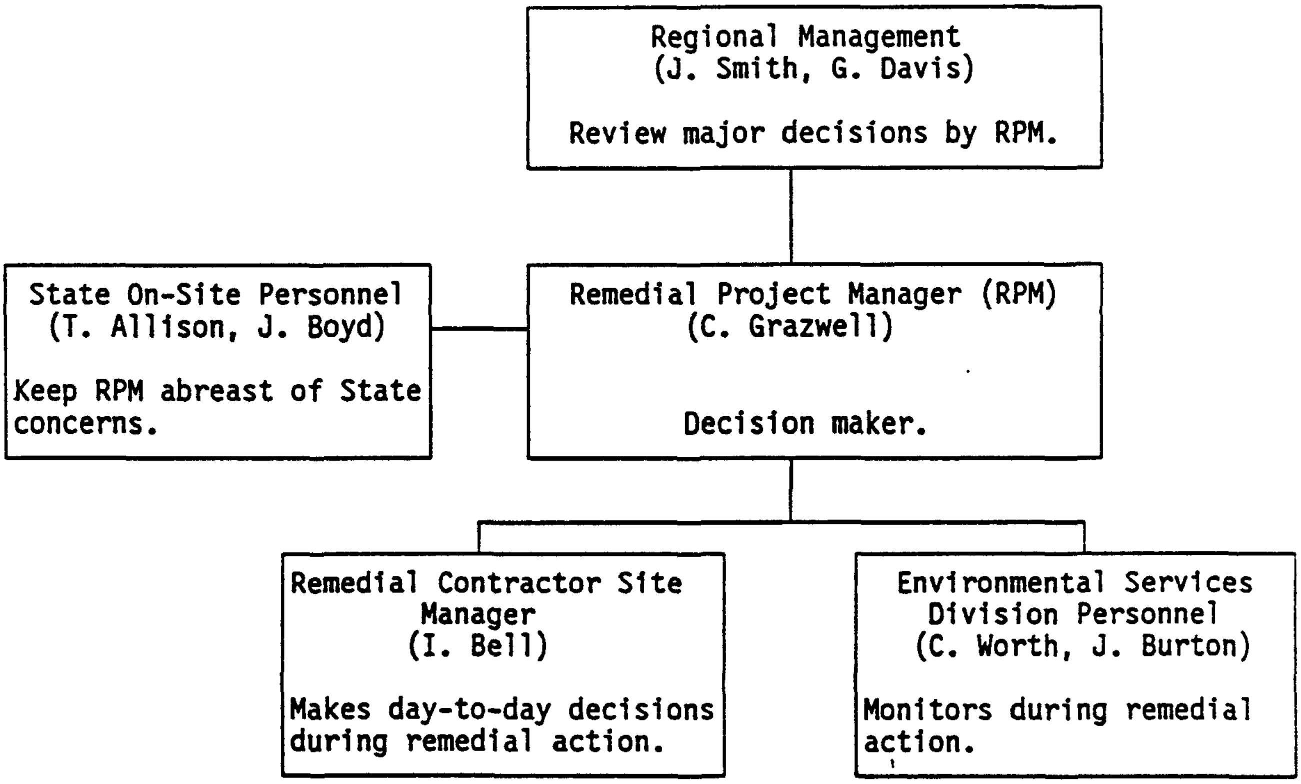
through use of overpacking, transport them to a nearby hazardous waste facility. This is expected to take 12 months of on-site activity, based on an 8-hour workday, five days a week.

Even though the main medium of concern at this site is groundwater, there is also concern for the potential release of VOCs to ambient air. During the RI/FS, drums were disturbed and an unusual odor was detected at one of the neighboring houses. An FID measurement indicated that VOCs had been released into the air. The monitoring will be necessary to detect any subsequent releases to the ambient air which would pose a threat to public health during remediation. The guidance given in Procedures for Conducting Air Pathway Analyses for Superfund Applications will be followed in designing the ambient air quality monitoring system for this site. If the ambient air concentration level of concern of one or more of the VOCs is exceeded, the decision maker for the remediation activity must implement the appropriate action for reducing the concentration or evacuating the subject population.

There are six houses between 100 to 200 meters of the site. The safety and health of the residents of these houses and the on-site workers during remediation must be assured. For example, the use of protective clothing for the on-site workers insures their safety, and the implementation of an air quality monitoring system that will detect with high probability elevated levels of air pollutants in a timely manner at the receptor sites insures the safety of the subject population.

2.2 DECISION MAKERS AND KEY DATA USERS

The organizational chart that follows presents the relationships between the organizations involved in Superdump during remedial action. Names of individuals and their functions for each organization are given. The Regional Management has final authority on the site cleanup and, as such, is the ultimate decision maker in terms of setting the acceptable levels of uncertainty in answering the question, "Has the remedial response activity resulted in pollutant concentrations above the level of concern at receptor sites?" The RPM is the on-site decision maker in terms of reviewing the monitoring data and making the decision to either change the monitoring strategy, implement emission control procedures, stop the remedial response activity, or evacuate the population. State on-site personnel are to observe the activities for compliance with agreed-upon procedures and to further insure that the subject population is properly protected. The Remedial Contractor Site Manager will coordinate daily activity at the site. Environmental Services Division personnel will monitor during remedial response activities. The RPM is the key data user and decision maker for this site.



2.3 DECISION DEFINITION

The main decision to be made (and the actions to be taken or not taken) depends on the answer to the question, "Has a remedial response activity at this site caused ambient air concentrations of VOCs to exceed the acceptable risk-based concentrations or applicable or relevant and appropriate requirements (ARARs)?" If ARARs or the other concentration levels of concern are exceeded, the decision maker must decide what action to take to protect the public's health. The action may be:

- to institute controls to lower air emissions,
- to halt remedial activities, and/or
- to evacuate the receptor population.

If the concentration levels are below the levels of concern, the decision maker will elect to continue with the remedial response activity. Also, in instances where the measured pollutant concentration increases or decreases, but remains below the level of concern, the decision maker's action may be to change the monitoring strategy to one with a higher probability of detecting an exceedance in the first case or to a more relaxed strategy in the second case.

2.4 AVAILABLE RESOURCES

The available budget for air monitoring during remedial action at Superdump has been tentatively limited to $50,000. If adequate monitoring cannot be performed for this cost, negotiation for more funds will be conducted. If additional funds are not available through negotiation, it may be necessary to proceed with a monitoring system that could lead to a greater risk of making an incorrect decision.

Various applicable sampling, analysis, and modeling methods are available. A portable flame ionization detector (FID) is available for instances where measurement results are needed immediately and when the level of concern that will prompt action is in terms of a total hydrocarbon concentration instead of a specific compound. If the decision requires compound-specific concentrations, gas chromatography (GC), gas chromatography/mass spectrometry (GC/MS), or high performance liquid chromatography (HPLC) may be chosen during the design of the monitoring system. If quick turnaround is important, portable and transportable GCs are available. In an emergency situation where pollutant levels are suspected of being greatly above the levels of concern, a quick and inexpensive method like a colorimetric detection tube may be adequate. For example, GC/MS air analysis with 48-hour turnaround could cost five times what a 2-week turnaround would cost.

For Superdump, Environmental Services Division (ESD) personnel have access to a mobile laboratory equipped with a generator and containing GCs with a variety of columns and detectors and portable FIDs. If appropriate, prior to the start of remedial activities, colorimetric detector tubes can be purchased. Also, arrangements with an off-site laboratory for sample analysis by GC/MS can be made if deemed necessary. GC/MS analysis at the nearest off-site laboratory has a 48-hour turnaround time. Real-time meteorological data are available from the State air monitoring station.

Modeling may be used to logically determine the placement of monitoring stations around Superdump and to calculate worst-case exposures at the receptor sites based on analytical measurements taken either where the digging occurs or at the fenceline of the site. Different levels of sophistication exist for modeling as well as for sampling and analysis. These levels range from screening models, which assume that the contamination plume will travel directly to the receptor, to refined models that consider other factors, such as meteorological conditions and variations in source strength. Modeling combined with results from sampling and analysis may be used to characterize the site and surrounding area.

Two technicians are available for operation of the ambient air monitoring equipment. Additional equipment and manpower can be made available if needed, but only at increased cost. Depending on the magnitude of additional equipment and manpower needed, the budget may need to be increased from two to five times. All of these details will be resolved in Stage III of the DQO process when the monitoring system is designed. For now, it is necessary only to gather information on available methodologies and instrumentation for analysis of the pollutants of interest. If an acceptable analytical method is not available for a pollutant of interest, method development may be necessary or a surrogate measurement, such as total VOCs, may be identified for monitoring.

2.5 INFORMATION REQUIREMENTS

At this point, the definition of required information need not be stated in the form of the variables to be measured (e.g., particulate matter or VOC concentrations), but in more general terms, such as information on potential receptors near the site, concentrations of air toxics, ARARs, and health effect information on specific pollutants at the site. It is appropriate to begin to consider what specific aspect of ambient air quality at the Superfund site will bring about any particular action and what areas of ground surface and intervals of time are of interest. Types of information needed to support the design of the monitoring system are:

- Information on specific pollutants present at the site (their identity, concentration range, health effects, and concentration levels of concern). As mentioned in section 2.1, the drums at this site contain mainly solvents. The RCRA waste codes are F001, F002, F003, F004, F005, and F006. The chemicals in these wastes and their air toxicity constants are listed in Table 4 in section 3.2.

- Data from monitoring of the undisturbed site (meteorological and ambient air quality data). During the RI/FS, an FID measurement indicated that VOCs were released when the drums were disturbed. No compound-specific measurements were made at that time. The average wind speed for this site is 2 meters/second and the prevailing direction is from the northeast quadrant in the winter and from the southwest quadrant in summer.

- Information on potential receptors near the site (population distribution and livestock, for instance). The potential receptors at this site are the residents of the six houses.

- Political, social, and economic issues. After the unusual odor was noticed at a neighboring house during the RI/FS, several neighborhood meetings were held to discuss activities at the site. The concerns of the neighbors will need to be addressed during the remedial response activity.

- Removal response activity that is planned. The remedial action for this site is to dig up the buried drums, overpack them, and transport them to a nearby hazardous waste facility. The monitoring system will need to consider possible releases of VOCs to the ambient air during each of these steps. This type of activity will almost always be covered by an on-site coordinator and not an RPM.

- Regulatory and legal concerns. The ambient air quality must be in compliance with all local, State, and Federal regulations.

Of the above bulleted information, only the first, information on specific pollutants present at the site, will require the collection of new environmental data.

2.6 DATA REQUIREMENTS AND USES

For monitoring at Superfund sites, real-time (or as near real-time as possible) data are required. Existing data from a different temporal domain may be helpful for designing a monitoring program, but cannot be used to make real-time decisions. The decision maker will also need to state how quickly measurement data results are needed for actions to be taken to protect the public health in case of an exceedance.

Ambient air data from monitoring during the remedial response activity will be used to answer two major questions:

1. What are the ambient air VOC concentrations?

2. Are there any levels of concern that are exceeded at these concentrations?

Monitoring data collected during the remedial response activity which indicate that ambient air concentrations of concern are exceeded may require one or more of the following actions:

- management practices to control air emissions to an acceptable level,

- scheduling of the remedial operation for a day when the wind is forecast to be blowing away from the houses of concern,
- the remedial activity to be halted and the affected population evacuated, or
- the population to be evacuated until the remedial activity is completed.

Management practices could consist of such extreme actions as building a shed over the site with a vacuum system that collects the air in the shed and pumps it through an activated charcoal system before releasing it to the ambient air. Where evaporation from soil contaminated by leakage from the drum is the primary problem, a less costly action would be to spray the ground with foam for short term relief. If form dissipates during strong winds, for example, reapplication could be costly. If preliminary monitoring indicates that there may be a problem (i.e., concentrations near the ARARs or other levels of concern), more frequent or more intense monitoring (more samplers or different types of monitors) may be deployed for the duration of the remedial action to increase the probability of detecting an exceedance.

If the monitoring data indicate that the remedial response activity has not created an ambient air problem, the activity will proceed, provided there are no problems in other media.

2.7 CONSEQUENCES OF MAKING INCORRECT DECISIONS

Incorrect decisions include unnecessarily taking action to reduce pollution concentrations (a false positive error) or not taking action when an action should be taken (a false negative error). The types of decision errors are illustrated in Table 3. For example, if measurement data indicate that one or more pollutants exceed the concentration levels of concern when in fact they do not, unnecessary costs associated with taking action to control air emissions may be incurred. There may also be unnecessary costs and inconveniences associated with evacuating the affected population, and increasing the level of monitoring (either by using more monitors or a more sophisticated monitoring technique) for the duration of the remedial action. This type of error, deciding to take unnecessary action, is a false positive error. A more detailed discussion of false positive and false negative errors is given in Appendix A.

TABLE 3. DECISION ERRORS

	Actual Concentration Less than Level of Concern	Actual Concentration Equal to or Greater Than Level of Concern
Take No Action	No Error	False Negative Error
Take Action	False Positive Error	No Error

If measurement data indicate that pollutants do not exceed the levels of concern when in fact they do, and remedial activity continues with no action taken to reduce emissions, there may be detrimental health impacts. This type of error, not taking action when action is needed, is a false negative error. The consequences of a false negative error depends on the seriousness of the undetected problem. If, for example, the concentration is in the range associated with possible health risk and it is not detected, this is serious. If the concentration is at the level which poses an immediate acute health risk and is not detected, this is very serious. This situation should be detected by the monitoring system with almost certainty.

The ranking of errors by their seriousness for Superdump is as follows:

ERROR	LEVEL OF SERIOUSNESS	ACTION
• False Positive: Concluding there is a problem when the concentration is well below a level which poses a health risk.	Least Serious (Resources used unnecessarily)	Halting remedial response activity or instituting management practices to control air emissions.
• False Negative: Concluding there is no problem when the concentration is in the range associated with health risk.	Serious (Possible public health problem)	Taking no action to reduce/control emissions or to protect the population.
• False Negative: Concluding there is no problem when the concentration is at or near a level which poses an immediate acute risk to health.	Very Serious (Likely public health problem)	Taking no action to reduce/control emissions or to protect the population.

SECTION 3.0

STAGE II: REFINEMENT OF THE DECISION AND REQUIREMENTS

In Stage I, a list of actions for protecting the population through control of emissions or evacuation has been established, together with an initial description of the types of environmental information needed to prompt the decision maker to take these actions. The goal of Stage II is to clarify these initial statements and define what types of measurement data need to be collected over what areas of space and intervals of time.

3.1 QUESTIONS TO BE ANSWERED

The decision of whether or not to take an action will be based on the answer to the following question:

> To what extent are remedial activities at Superdump creating an ambient air pollution problem?

This question can be broken down into simpler, manageable questions which can be answered using monitoring data:

1. Has the concentration level of concern, from an ARAR, health risk-based calculation, or other criteria, been exceeded in areas of concern by one or more VOCs (nearby off-site areas where receptors are located)?

2. If the concentration level of concern has been exceeded in an area of concern, was the exceedance due entirely to activities at the site? What portion of the exceedance was due to background or other pollution sources?

The ambient air monitoring data needed to answer the first question are the air quality measurement data at the receptor sites or air quality measurements at the fenceline combined with meteorological measurements and/or model predictions made for the receptor locations. The second set of questions need to be answered only if the concentration level of concern is exceeded at any of the off-site receptor locations. Environmental data needed to answer this second set of questions are:

- Meteorological data on wind speed and direction during the period in which the concentration level of concern was exceeded at a specific receptor site

- Results of downwind fenceline and/or off-site monitoring
- Results of monitoring upwind of or adjacent to the site and at the active area of the site
- Any anecdotal information on remedial response activity at Superdump which would explain an increase in downwind concentration levels

If the level of concern is exceeded at a receptor site, the decision maker needs to answer the second set of questions before he or she can determine if on-site emission control actions will appreciably reduce the concentration at the receptor location or if the only recourse is to halt the remedial response activity to assure the population that Superdump is not the cause of the problem.

3.2 VARIABLES AND DOMAIN

For this Superfund site, the receptor population that might be exposed to high levels of volatile organics in ambient air are the residents of the nearby houses. These six houses are within about 100 meters of Superdump and constitute the spatial domain of concern for decision making. For this example, the workers on-site wear respirators and protective clothing adequate to protect them from excessive volatile organics in ambient air. The temporal domain is the 8-hour workday, five days a week, for the 12 months of the remedial response activity.

Since site records indicate that RCRA waste codes F001, F002, F003, F004, F005, and F006 are contained in the drums, information on the chemicals in these wastes (from the RCRA waste code description) and their properties (from Appendix A of the Superfund Public Health Evaluation Manual) is needed. This chemical-specific information is summarized in Table 4. The information was reviewed, and the top six chemicals based on vapor pressure, Henry's Law constant, and air toxicity constant were selected as indicators. These top six chemicals are a subset of the indicator chemicals selected during the baseline risk assessment during the remedial investigation phase. These indicator chemicals were judged to be the most toxic and most volatile ones at the site and are listed in Table 5.

TABLE 4. CHARACTERISTICS OF CHEMICALS AT THE SITE

RCRA CODE	CHEMICAL NAME	CAS NUMBER	VAPOR PRESSURE (mm Hg)[A]	HENRY'S LAW CONSTANT (atm m^3/mol)[A]	AIR TOXICITY CONSTANT (carcinogenic) (m^3/mg)[B]	AIR TOXICITY CONSTANT (non-carcinogenic) (m^3/mg)[C]
F001, F002	Tetrachloroethylene	127-18-4	1.78+01	2.59-02	8.86-02	2.75-02
	Trichloroethylene	79-01-6	5.79+01	9.10-03	4.29-02	2.96+01
	Methylene chloride					
	1,1,1-Trichloroethane	71-55-6	1.23+02	1.44-02	---	7.33-03
	Carbon tetrachloride	56-23-5	9.00+01	2.41-02	1.88-01	3.17+00
	Chlorobenzene	108-90-7	1.17+01	3.72-03	---	2.79-01
	1,1,2-Trichloro-1,2,2-Trifluoroethane	76-13-1	2.70+02	---	---	---
	o-Dichlorobenzene	95-50-1	1.00+00	1.93-03	---	3.61-01
	Trichlorofluoromethane	75-69-4	6.67+02	---	---	---
	1,1,2-Trichloroethane	79-00-5	3.00+01	1.17-03	1.03-01	---
F003	Xylenes	---	1.00+01	---	---	---
	Acetone	67-64-1	2.70+02	2.06-05	---	---
	Ethyl acetate	141-78-6	---	---	---	---
	Ethyl benzene	100-41-4	7.00+00	8.43-03	---	1.10-01
	Ethyl ether	---	---	---	---	---
	Methyl isobutyl ketone	108-10-1	---	---	---	---
	n-Butyl alcohol	71-36-3	---	---	---	---
	Cyclohexanone	---	---	---	---	---
	Methanol	67-56-1	---	---	---	---
F004	Cresols & cresylic acid	1319-77-3	2.40-01	1.10-06	---	5.97+01
	Nitrobenzene	98-95-3	1.50-01	---	---	---
F005	Toluene	108-88-3	2.81+01	6.37-03	---	5.20-02
	Methyl ethyl ketone	78-93-3	7.75+01	2.74-05	---	7.75-02
	Carbon disulfide	75-15-0	3.60+02	1.23-02	---	4.24+00
	Isobutanol	78-83-1	---	---	---	---
	Pyridine	110-86-1	2.00+01	---	---	---
	Benzene	71-43-2	9.52+01	5.59-03	7.71-02	1.18+02
	2-Ethoxy ethanol	---	---	---	---	---
	2-Nitropropane	---	---	---	---	---
F006	Cadmium	7740-43-9	0	---	1.65+01	3.59+02
	Chromium^{+6}	7740-47-3	0	---	1.11+02	2.50+01
	Chromium - total	7740-47-3	0	---	---	---
	Nickel	7440-02-0	0	---	2.85+00	1.57+02
	Cyanide	57-12-5	Varies	---	---	---

NOTE: Values are in scientific notation.

[A] From Appendix A, Exhibit A-1 of EPA's Superfund Public Health Evaluation Manual.
[B] From Appendix A, Exhibit A-3, Toxicity Data for Potential Carcinogenic Effects.
[C] From Appendix A, Exhibit A-5, Toxicity Data for Noncarcinogenic Effects (Note: Toxicity constants for carcinogens should not be compared with toxicity constants for noncarcinogens).
----indicates missing data.

TABLE 5. EXPOSURE LIMITS FOR SELECTED INDICATOR CHEMICALS

Chemical	Permissible Exposure Limit (PEL)	Short-Term Exposure Limit (STEL)
Benzene	1 ppm for 8-hour time weighted average	5 ppm for 15 minutes[A]
Trichloroethylene	50 ppm for 8-hour time weighted average	200 ppm for 15 minutes[B]
Carbon tetrachloride	2 ppm for 8-hour time weighted average	------
1,1,2-Trichloroethane	10 ppm for 8-hour time weighted average	------
Carbon disulfide	4 ppm for 8-hour time weighted average	12 ppm for 15 minutes[B]
Methyl ethyl ketone	200 ppm for 8-hour time weighted average	300 ppm for 15 minutes[B]

ppm = parts per million

[A] 29 Code of Federal Regulations 1910.1028, Subpart Z.

[B] Federal Register, 54, Thursday, January 19, 1989, pages 2332-2983.

----indicates no short-term exposure limit.

3.3 SUMMARY RESULTS

The summary results to be used in decision making depend on how the level of concern is expressed. For instance, if the level of concern is expressed as an 8-hour average, 8-hour averages will be used for decision making. Information on permissible exposure limits (PELs) and short-term exposure limits (STELs) for the six indicator chemicals are also presented in Table 5. There are no ARARs for these chemicals at this site. Toxicity information from the Superfund Public Health Evaluation Manual is not applicable to this scenario because it is for chronic exposure, whereas during remediation, the concern is for short-term exposure. If the excavation site is left open after the 8:00 a.m. shift, vapors could still be released. In that event, TLVs and STELs would not be appropriate. The exposure limits in Table 5 are from the Occupational Safety and Health Administration (OSHA) regulations. Since these regulations are designed for protecting workers' safety and health on an 8-hour basis, they do not represent the decision criterion for this site, except for the on-site workers. Workers are healthy adults, in general, but the population near Superdump may include children (with lower body weights), elderly, and sick individuals. To take into account the more sensitive population and to err on the side of safety as well, it was decided to use 1/10 of the 8-hour PELs for making daily decisions. In addition, if an STEL is reached, immediate action will be necessary since these are 15-minute exposure limits for worker safety and health. Thus, there are two allowable exposure limit summary statistics for each of the selected indicator chemicals: 8-hour time weighted averages and 15-minute averages.

3.4 NEED FOR NEW DATA

All of the environmental data identified as needed to answer the questions in section 3.1 have to be new data to represent the temporal domain of interest. That is, since it is expected that the emission rate from Superdump will vary with time, and thus, ambient air pollutant concentrations will vary in time and space, real-time (or near real-time) monitoring data will be needed by the decision maker to effectively protect the subject population during the remedial response activity.

3.5 ACCEPTABLE ERROR RATES

For each of the six compounds, there are two levels of concern at the receptors: the short-term exposure limit and 1/10 of the permissible exposure limit. Based on consideration of the consequences of making decision errors, the RPM established the following as acceptable error rates for this site (Note: general statements of acceptable error rates are followed by specific examples using benzene):

- At a true average concentration of 1/2 of either concentration level of concern, the probability of a positive finding should be limited to less than 10%; that is, at least 90% of the time, the data would correctly indicate that there is no problem.

  For example, if the true benzene concentration is 2.5 ppm for a 15-minute average, there should be no more than a 10% chance of the monitoring/modeling information indicating a concentration of 5 ppm or more (that is, the probability of a false positive error is 10% or less).

- When the true average concentration is 1 1/2 times either level of concern, the probability of negative findings should be limited to less than 5%; that is, at least 95% of the time, the monitoring data would correctly indicate that there is a problem.

  For example, if the true benzene concentration is 150 ppb for an 8-hour average, there should be no more than a 5% chance of obtaining a determination of 100 ppb or less (that is, the probability of a false negative error is 5% or less).

- When the true average concentration is 2 times either level of concern, the probability of a negative finding should be limited to less than 1%; that is, at least 99% of the time, the monitoring data would correctly indicate that there is a problem.

  For instance, if the true benzene concentration is 200 ppb for an 8-hour average, there should be no more than a 1% chance of obtaining a determination of 100 ppb or less (that is, the probability of a false negative error is 1% or less).

Based on the above error rates as specified by the decision maker and the exposure limits given in Table 5, compound-specific DQOs are given in Table 6.

Figure 2 is a graphical representation of the DQOs from Table 6 and, as described above, for short-term exposure to benzene which has a STEL of 5 ppm. The straight lines forming rectangles show the DQOs while the S-shaped curve shows how one sampling and analysis strategy is expected to perform. The graph shows that for actual concentrations between 0 and 2.5 ppm (1/2 of the STEL), the probability of taking action (the probability of a false positive)

TABLE 6. COMPOUND-SPECIFIC DATA QUALITY OBJECTIVES

Compound	Type of Exposure Limit	Concentration (1/2 level of concern)	Acceptable False Positive Probability	Concentration (1 1/2 x level of concern)	Acceptable False Negative Probability	Concentration (2 x level of concern)	Acceptable False Negative Probability
Benzene	Short-term	2.5 ppm	10%	7.5 ppm	5%	10 ppm	1%
	8-hour Permissible	50 ppb	10%	150 ppb	5%	200 ppb	1%
Trichloroethylene	Short-term	100 ppm	10%	300 ppm	5%	400 ppm	1%
	8-hour Permissible	2.5 ppm	10%	7.5 ppm	5%	10 ppm	1%
Carbon tetrachloride	Short-term	---	---	---	---	---	---
	8-hour Permissible	100 ppb	10%	300 ppb	5%	400 ppb	1%
1,1,2-Trichloroethane	Short-term	---	---	---	---	---	---
	8-hour Permissible	500 ppb	10%	1.5 ppm	5%	2 ppm	1%
Carbon disulfide	Short-term	6 ppm	10%	18 ppm	5%	24 ppm	1%
	8-hour Permissible	200 ppb	10%	600 ppb	5%	800 ppb	1%
Methyl ethyl ketone	Short-term	150 ppm	10%	450 ppm	5%	600 ppm	1%
	8-hour Permissible	10 ppm	10%	30 ppm	5%	40 ppm	1%

NOTE: Short-term exposure limits (STELs) are for 15-minute averages.
Permissible exposure limits (PELs) are for 8-hour averages.
For this site, the levels of concern are the STEL and 1/10 of the PELs (see section 3.3).
--- indicates no short-term exposure limit; thus, short-term DQOs could not be calculated.

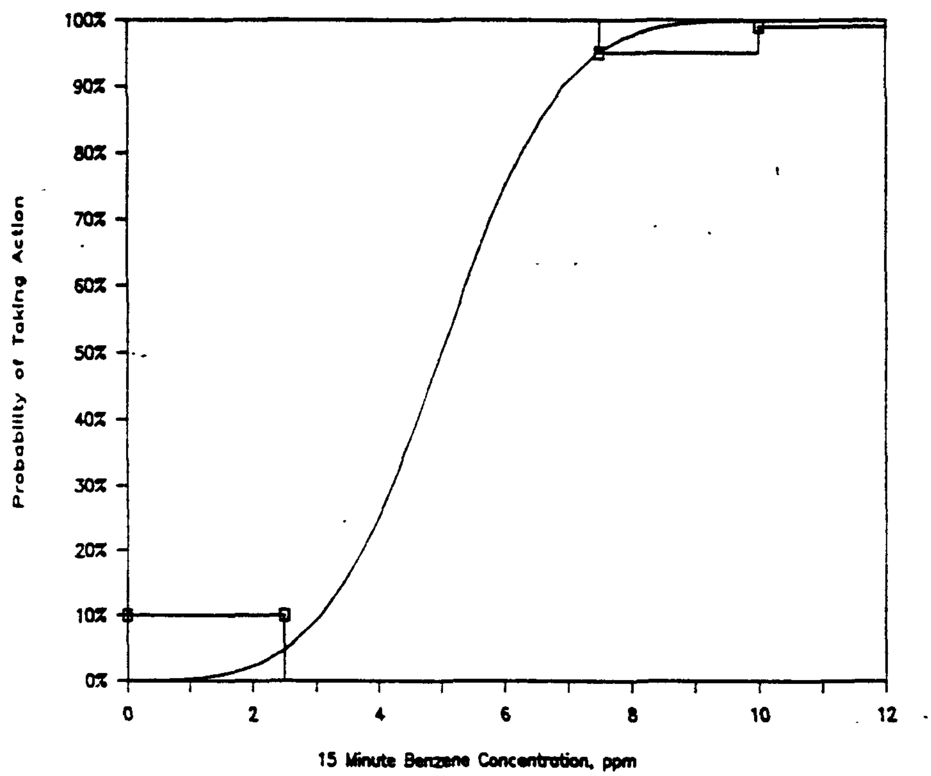


Figure 2. Short-term exposure DQOs for benzene at Superdump.

should be 10% or less. Between 7.5 (1 1/2 times the STEL) and 10 ppm (2 times the STEL), the probability of taking action (the probability of a true positive) should be 95% or more (thus, the probability of a false negative error is 5% or less). Finally, above 10 ppm, the probability of taking action should be 99% or more (thus, the probability of a false negative is 1% or less). The performance curve characterizes a sampling and analysis strategy which achieves or surpasses these DQOs across the entire range of concentrations.

In Stage III of the DQO process, a number of alternative monitoring/ modeling strategies will be considered, producing graphs similar to Figure 2. Acceptable strategies are only those which are expected to achieve the DQOs and at the same time do not exceed resource limitations.

3.6 SUMMARY OF THE DATA QUALITY OBJECTIVES

DEFINE THE DECISION

- <u>description of decision</u>
  The decision is whether action must be taken to protect the public from VOCs in ambient air during remedial response activities.

- <u>background for the problem</u>
  The site contains 2,000 buried drums with RCRA waste codes of F001, F002, F003, F004, F005, and F006. The remedial response activity will be to dig up the drums, overpack them, and transport them to a hazardous waste facility.

- <u>actions under consideration if a problem is found</u>
  Any of the following actions may be considered singularly or in combination with other actions:
  - institute controls to lower air emissions
  - halt remedial activity
  - evacuate the receptor population which is at risk
  - employ a more rigorous monitoring strategy

- <u>domain of the decision</u>

  The spatial domain for this decision is the ambient air at the receptor locations (six houses) and the ambient air within the site boundaries (for on-site workers). The temporal domain is the 8-hour workday, five days a week, for the year of the remedial response activity. Unless vapors from excavation are controlled when work is not in progress, the above scenario is inappropriate. To make sure vapors are not being emitted during non-working hours, monitoring may be required.

- <u>information needs</u>

  The decision will be based on the determination of short-term (15-minute) and 8-hour average concentrations of contaminants in ambient air. The contaminants of concern and their exposure limits will be selected based on information on toxicity, the waste types expected, and potentially sensitive receptors.

DEFINE THE USE OF DATA

- <u>elements of decision which require data</u>

  The elements for Superdump are the compounds of interest. They are benzene; trichloroethylene; carbon tetrachloride; 1,1,2-trichloroethane; carbon disulfide; and methyl ethyl ketone concentrations in ambient air. These elements are quantifiable and data-dependent. The temporal domain is the entire time the site is active (normally 8 hours per day, 5 days per week) broken into 15-minute segments for short-term exposure measurements.

- <u>use of data-dependent elements</u>

  If measurement data show that the concentration of any one contaminant of concern exceeds the STEL, then action must be taken to reduce the VOC emissions. In addition, more rigorous monitoring will be required to determine the extent of any off-site exceedance of the STEL. Such an exceedance may be cause for evacuation.

  If the highest average 8-hour exposure at any receptor is above the level of concern (1/10 the PEL), then action must be taken to reduce the VOC emissions.

Since there are usually so many compounds, the levels of all compounds added together may cause adverse health effects. ACGIH uses the following formula to determine action levels for on-site personnel.

If $C_1/TLV_1 + C_2/TLV_2 + \ldots C_n/TLV_n \geq 1$
(where C_n is the concentration of the nth component and TLV_n is the threshold limit value of the nth component), action is required. If level < 1, then no action is taken.

RANKING OF ERRORS ACCORDING TO SERIOUSNESS OF THEIR CONSEQUENCES

- <u>The ranking of decision errors by their seriousness is as follows:</u>

  -- Concluding there is a problem when the concentration is well below the level which poses a health risk is the least serious error.

  -- Concluding there is no problem when the concentration is in the range associated with a health risk is serious.

  -- Concluding there is no problem when the concentration is in the range associated with immediate acute health problems is a very serious error.

RESTATE THE DECISION

- <u>the decision</u>

  The decision is whether actions must be taken to protect the public from VOCs in ambient air during remedial response activities.

- <u>decision elements</u>

  Data needs: 15-minute and 8-hour average measurements of selected target contaminants in ambient air.

  Spatial domain: need to measure directly, or predict from measurements using models, the concentrations of contaminants at off-site areas where receptors are located. Data on background concentrations will be needed if concentration thresholds are found to be exceeded.

  Temporal domain: measurements will span the active period of each working day (typically 8 hours). When the site is inactive, concentrations are assumed to be less because the site will be left in a secure state. (No disturbed drums will be left exposed, and any spillage will be recovered prior to shutting down for the day, however, to substantiate hypothesis, monitoring may still be needed.)

FORMALIZE THE DECISION PROCESS

- <u>decision rule</u>

  If either exposure limit (STEL or 1/10 of PEL) is exceeded for any of the contaminants of concern, immediate action will be taken to control the release of volatiles.

- <u>role of data</u>

  The decision on whether or not to take action will be based on the ambient air measurement data.

- <u>need for new data</u>

  Since the site has not yet been remediated, data for the site during remediation are not available. All of the environmental data will need to be new data to represent the temporal domain of interest.

DESIRED PERFORMANCE - ACCEPTABLE PROBABILITY OF FALSE POSITIVE AND NEGATIVE ERRORS AT SELECTED CONCENTRATION LEVELS

- <u>false positive</u>

  At a true average concentration of 1/2 of either level of concern (STEL or 1/10 PEL), the probability of a positive finding should be limited to less than 10%.

  For example, if the true benzene concentration is 2.5 ppm for a 15-minute average, there should be no more than a 10% chance of obtaining a determination of 5 ppm or more.

- <u>false negatives</u>

  When the true average concentration is 1.5 times either level of concern, the probability of a negative finding should be limited to less than 5%.

  For example, if the true benzene concentration is 150 ppb for an 8-hour average, there should be no more than a 5% chance of obtaining a determination of 100 ppb or less.

When the true average concentration is 2 times either level of concern, the probability of a negative finding should be limited to less than 1%.

For example, if the real benzene concentration is 200 ppb for an 8-hour average, there should be no more than a 1% chance of obtaining a determination of 100 ppb or less.

INITIAL MONITORING SYSTEM DESIGN CONSIDERATIONS

- turnaround time
  Short-term measurements need to be available within 15 minutes. The 8-hour averages need to be available within 48 hours following the end of the sampling period.

- preventive maintenance / spare equipment
  Spare equipment should be available so that no period of time passes without monitoring activity. If equipment is not available and operational, remedial activity will not be allowed to proceed.

- limit of detection
  Measurement methods should have detection limits below 1/10 the concentration of concern for each compound. This applies to concentrations of concern for both short-term and 8-hour monitoring.

- monitoring flexibility
  Based on the actual compounds present and their concentration levels, there should be several monitoring strategies ranging from periodic screenings to on-site mobile or fixed monitoring stations with documented criteria for moving from one strategy to another.

- use of models
  The most cost-effective strategy may be a combination of monitoring and modeling.

- sampling error
  The spatial and temporal domains for this site must both be assumed to be heterogeneous. The VOC concentrations will vary with time and with location around the site. This heterogeneity may introduce the greatest amount of uncertainty to the decision making.

- <u>measurement errors</u>

  For the measurement methods discussed in section 2.4, the range of measurement error is from 1.1 to 2.0, expressed as uncertainty factors, according to Tables 3-4 and 3-5 in Volume IV of <u>Procedures for Conducting Air Pathway Analyses for Superfund Applications</u>.

REFERENCES

U.S. Environmental Protection Agency. Superfund Public Health Evaluation Manual. EPA/540/1-86/060, OSWER Directive 9285.4-1.

U.S. Environmental Protection Agency. Development of Data Quality Objectives, Description of Stages I and II. Quality Assurance Management Staff, July 16, 1986.

U.S. Environmental Protection Agency. Data Quality Objectives for Remedial Response Activities, Volume I, Development Process. EPA 540/G-87/003, March 1987.

U.S. Environmental Protection Agency. Data Quality Objectives for Remedial Response Activities, Volume II, RI/FS Activities at a Site with Contaminated Soils and Ground Water. EPA 540/G-87/004, March 1987.

U.S. Environmental Protection Agency. Data Quality Objectives for the Toxic Air Monitoring Program (Stages I and II). Office of Air Quality Planning and Standards, December 1987.

U.S. Environmental Protection Agency. Data Quality Objectives for the Urban Air Toxic Monitoring Program (Stages I and II). Office of Air Quality Planning and Standards, June 6, 1988.

U.S. Environmental Protection Agency. Procedures for Conducting Air Pathway Analyses for Superfund Applications, Volume I, Application of Air Pathway Analyses for Superfund, December 1988 draft.

U.S. Environmental Protection Agency. Procedures for Conducting Air Pathway Analyses for Superfund Applications, Volume II, Procedures for Developing Baseline Emissions from Landfills and Lagoons. EPA-450/1-89-002, January 1989.

U.S. Environmental Protection Agency. Procedures for Conducting Air Pathway Analyses for Superfund Applications, Volume III, Procedures for Estimating Air Emissions Impacts from Remedial Activities at NPL Sites. EPA-450/1-89-003, January 1989.

U.S. Environmental Protection Agency. Procedures for Conducting Air Pathway Analyses for Superfund Applications, Volume IV, Procedures for Dispersion Modeling and Air Monitoring for Superfund Air Pathway Analyses, December 1988 draft.

Data Quality Objectives Training Software Version 6.5, December 1988. Quality Assurance Management Staff, U.S. Environmental Protection Agency.

APPENDIX A

A DISCUSSION OF FALSE POSITIVE AND FALSE NEGATIVE ERRORS

Environmental decisions are often framed in terms of taking or not taking some action which is expected to help protect public health or the environment. An incorrect decision will usually result in either taking an action which is not warranted or not taking an action which is warranted. If taking an action is regarded as positive and not taking an action is negative, then the incorrect decisions are called false positives and false negatives, respectively. One consequence of a false positive, taking unnecessary action, could be that the pollution control process is made too strict and/or too costly. A consequence of a false negative, not taking the needed action, could be that the environment does not receive the protection it needs and, as a result, is either placed at risk or damaged.

For monitoring around Superfund sites, the greatest concern is almost always for the false negative because the consequences (people are placed at risk or the environment is damaged) are much more serious than those of the false positive (halting the remediation or increased cost of controls). The only case when there may be greater concern for the false positive would be where an emergency situation exists and unnecessarily halting remediation would have serious environmental and public health consequences.

The first step in evaluating potential decision errors is to list them in order according to the level of concern for their consequences. This ranking provides the initial basis for designing a monitoring program. The successful monitoring program design will provide adequate protection against each type of decision error.

Next, the technical staff and the decision maker must consider how erroneous experimental results could prompt the wrong decision to be made. For each type of error, an error rate will be assigned based on the following:

- Relative importance of the result (answer to the question) in making the decision, and
- Degree of concern for the type of decision error based on analysis of the consequences.

If a measurement result by itself prompts a decision, the probability of drawing the wrong conclusion must relate directly to the degree of concern for decision error. If the measurement answers but one of many questions which will together prompt the decision, then somewhat higher probabilities might be acceptable.

There are two ways in which the decision can be in error:

1. The null hypothesis is rejected, when in fact the null hypothesis is true (false positive or Type I error), or

2. The null hypothesis is accepted, when in fact the null hypothesis is untrue (false negative or Type II error).

Whenever a statistical test is done, the probability of the Type I error must be selected. As part of the testing process, a researcher chooses a significance level at which the hypothesis will be tested. If the test statistic, whatever it might be, falls outside this level, the null hypothesis will be rejected. If this occurs when the null hypothesis is true, it is known as a Type I error or as a "false positive." Commonly, the significance level or "alpha value" is chosen as 5% (or 1%). This means that if the test statistic falls outside the range of 95% (or 99%) of its assumed distribution, the null hypothesis is rejected.

Clearly, the possibility of a mistake is built into this system. It is given by definition that 5% (or 1%) of the time, the test statistic will fall outside this range when the null hypothesis is true. This probability must be accepted to test the hypothesis. The probability cannot be reduced to zero. Therefore, the probability of a Type I error (false positive) is set at a level which the investigator deems acceptable. The consequences of the Type I error are often economic ones involving the cost of unnecessarily halting the remediation or the cost of more stringent pollution control monitoring practices.

Concern must also focus on the other type of mistake, the "false negative" or Type II error. For monitoring at Superfund sites, these types of error are often more serious than the Type I error. For example, if monitoring data fail to detect a problem when a pollutant is actually present at levels which cause a serious health risk, then the decision maker would not know that some kind of protective action is needed. The consequence is that the population remains at risk.

In the case of false positives, the probability is set before any statistical test is chosen, or even before data are gathered. Unfortunately, the probability of occurrence of the Type II error is not so easily manipulated. This is because while there is only one null hypothesis, there are an infinite number of possible alternative hypotheses. In the false negative case, what is really sought is the probability that some or all of the alternative situations will be missed by the test.

Knowing the probability of missing all the possible alternative hypotheses is impossible. Therefore, the technique used is to pick one specific alternative hypothesis, or a set of them, which the researcher would not like to reject if they were true. For example, if a specific level of concern is exceeded by a negligible amount, say 1%, it may be permissible to miss it. If the level of concern is exceeded by a somewhat larger amount, the correct action should be taken with higher certainty. If the ARAR or other level of concern is exceeded by a very large amount (e.g., by a factor of 2), the decision maker should be very nearly certain of taking the correct action. One way to express these desires is in terms of power of the design. The power of a design is defined as 1 minus the probability of a Type II error. Thus, if a design has 95% power when the level of concern is exceeded by a factor of 2, then we would expect to correctly conclude there is a significant pollution problem 19 out of 20 times when the level of concern actually is exceeded by a factor of 2.

The relationship between Type I and Type II errors is graphically illustrated in Figure A-1 on two different scales. The graph has curves which depict two normal (i.e., Gaussian or "bell-shaped") distributions. The leftmost curve represents a set of environmental data obtained by measuring 1000 samples taken from a parcel of ambient air with a pollutant concentration of 100. Due to random error in sampling and analysis, there is a spread in the measured values. Most of the values fall near 100 with fewer values occurring as the distance from 100 gets larger. The dispersion of these values is the variability of the measurement error in our sample collection and analysis. The leftmost curve can be mathematically described by its mean, 100, and its standard deviation, 1.5. The rightmost curve has the same dispersion (standard deviation = 1.5), but a mean of 107. Suppose the null hypothesis to be tested is that the pollutant concentration is 100. We plan to take some kind of pollutant management action if the data show that the true concentration is greater than 100. The leftmost curve shows that measurement error can give us results greater than 100 even when the true concentration of the

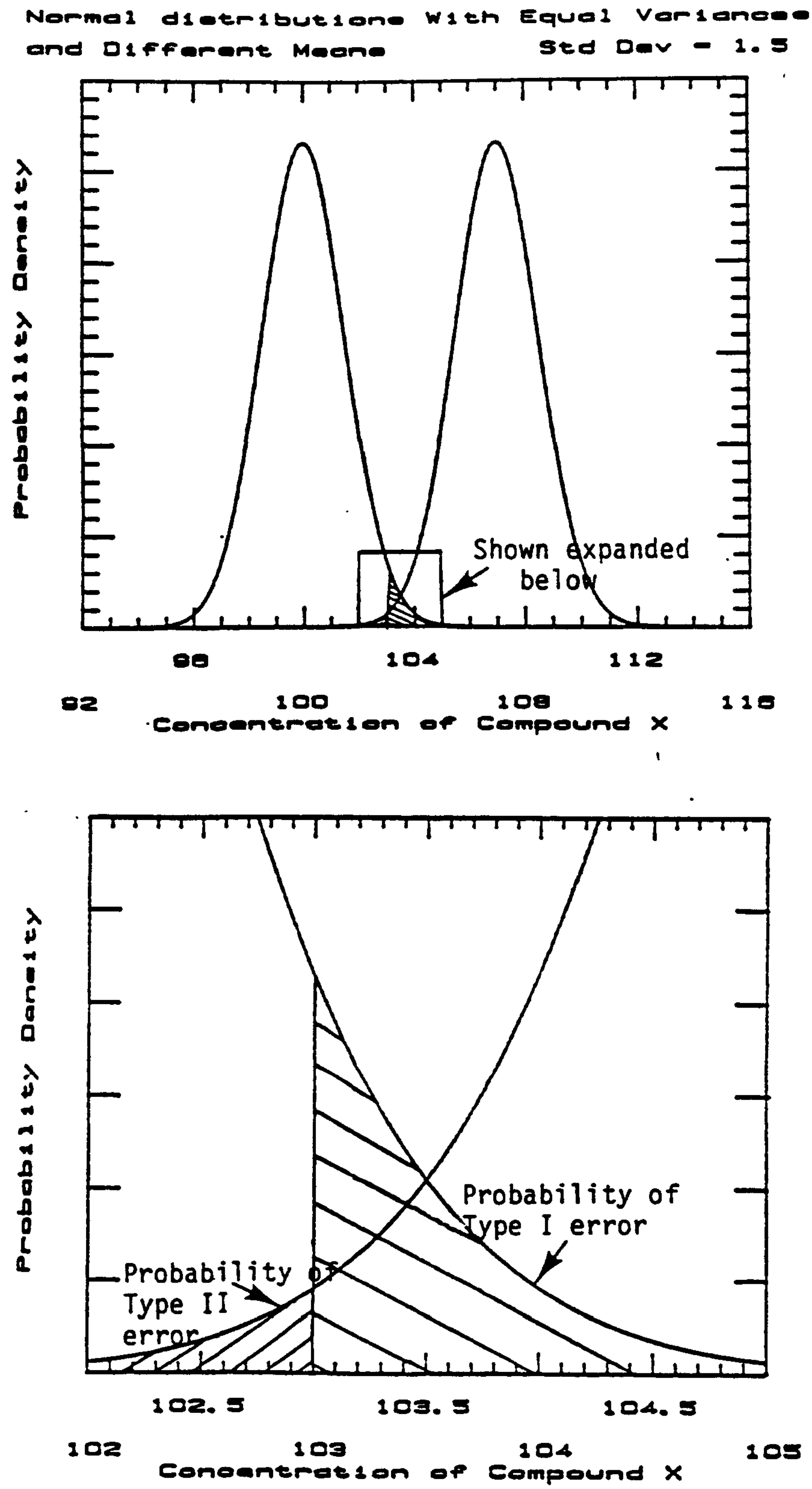


Figure A-1. Illustration of the relationship between Type I and Type II errors.

ambient air is exactly 100. But we would not want measurement error to cause us to take action when action is not needed, so a higher value, called an action threshold, is set so that measurement error has a small probability of causing action when the null hypothesis is true. There will be a 5% probability of causing this Type I error (false positive) if we set the threshold at 103. The probability of the Type I error is shown as the small shaded area to the right of 103 above the graph's X axis. A smaller probability for the Type I error can be achieved by moving the action threshold further to the right or by improving the precision of the measurements.

Now, suppose the true situation is that the mean is equal to 107. This produces the rightmost curve in Figure A-1. The very small shaded area to the left of 103 shows the area of the true distribution which falls below the critical value of X=103. These values of the test statistic will not lead to a rejection of the null hypothesis, even though it should be rejected. This shaded area is the probability of a Type II error. The remaining area under the righthand curve is 1 minus the probability of a Type II error, or the power of the test.

Usually the probability of a Type I error is prescribed and is independent of the experimental design, but the Type II error can only be determined after the conditions of the experiment are clearly defined. Thus, a specific statistical test must be chosen, with a certain number of samples, a given null hypothesis, a given alternative hypothesis, and certain parameters of the population under study (such as means and variances). These items will be important in Stage III, when the monitoring system is designed and assessed. At this point, it is sufficient to remember that for any given experimental design and alternative hypothesis, a certain power exists, and that by changing certain conditions of the experiment (e.g., sample size or significance level of the test), power can be adjusted to achieve the desired level, although the cost of the experiment may be intolerable.

APPENDIX B
GLOSSARY

GLOSSARY

ARARs - Applicable or relevant and appropriate requirements.

Domain - The spatial and temporal definition of the environment which is subject to the decision and action. The purpose of sampling and analysis is to characterize the domain.

DQIs - Data quality indicators. These include precision, accuracy, representativeness, completeness, comparability, and method detection limits.

DQOs - Data quality objectives. DQOs are statements of the level of uncertainty which the decision maker is willing to accept in the results derived from environmental measurements. For projects which are planned to test specific hypotheses (such as whether a pollutant concentration exceeds a standard), this uncertainty is expressed in terms of the probability of making Type I and Type II errors. (Type I error is the rejection of a null hypothesis which is actually true, and Type II error is the failure to reject a null hypothesis which is actually false.) For projects which are to produce interval estimates, this uncertainty is expressed as a desired confidence interval width.

EPA - U.S. Environmental Protection Agency.

EPMs - Enforcement Project Managers.

FIDs - Flame ionization detectors.

FS - Feasibility study.

GC - Gas chromatography.

GC/MS - Gas chromatography/mass spectrometry.

HPLC - High performance liquid chromatography.

Hypothesis - A hypothesis is a statement about a particular population parameter. Spatial areas and temporal intervals to be included must be defined. Some value for comparison must be included. A hypothesis is tested by measuring the actual value of the parameter (the statistic) in the system of interest, and computing the probability of this measured value occurring if the hypothesis is true. If the probability of observing the measured statistic is very low, given that the hypothesis is true, then the hypothesis is rejected.

Questions are generated to refine or test a model of the system under study. In order to be tested statistically, general questions about the model must be reduced to hypotheses that can be stated in terms of a specific statistic, which is some function of the collected data. Each statistical test requires a null hypothesis. Given the null hypothesis, the statistic chosen is then examined to see if its value is probable or improbable, assuming the hypothesis is true. If the value is probable, the null hypothesis is accepted. If the value is improbable, the null hypothesis is rejected, and some other model is assumed to describe the population. If our model is the simplest one (for example, that the variable of interest is normally distributed), then the null hypothesis tested would concern the mean and variance and how well the data fit a normal distribution. If the model involves relationships between variables, a testable null hypothesis is that the correlation coefficient is equal to 0.

The goal of the planning process is to develop one or more such statistical hypotheses for which data will be collected to test. Based on the results of the hypothesis tests or parameter estimations, decisions will be made. The design of a data collection program that most efficiently supports the decision requires an early statement of these hypotheses.

Model - A description of the interrelationship between variables in the particular system being studied. It describes, in greater or lesser detail, how the dependent variable(s) react to changes in the independent variables. This model, if it is to work well and lead to increased understanding, should reflect the cause-and-effect relationships among the variables. In the early phases of research, however, models may only show correlations, not necessarily causes.

Observation - Also called a data point, a measurement, or a value. The number derived from a measurement of a particular variable at a certain point in space and time. A data base is a set of observations of one or more variables over space and/or time.

OSHA - Occupational Safety and Health Administration.

Parameter - A number derived from a set of observations according to some rule. The number may correspond to an actual observation or to a mathematical function which combines the observations in some way. The term "parameter" can be used to refer to the function itself, such as the mean or the maximum, or to the unknowable, "true" value of the parameter as estimated from a particular set of observations (see statistic).

PARCC - Precision, accuracy, representativeness, completeness, and comparability.

PEL - Permissible exposure limit.

PIDs - Photoionization detectors.

Power - The probability of detecting a departure from the null hypothesis. When the null hypothesis is true, power is the probability of incorrectly rejecting it. This is a Type I error. When the null hypothesis is false, the term "1 minus power" is the probability of a false negative. This is a Type II error.

QA - Quality assurance. A system of activities whose purpose is to provide adequate confidence that a product or service will satisfy given needs.

QAPjP - Quality assurance project plan.

QC - Quality control. The operational techniques and the activities which are aimed at maintaining a product or service at a level of quality that will satisfy given needs.

RCRA - Resource Conservation and Recovery Act.

RI - Remedial investigation.

RPMs - Remedial Project Managers.

Significance - One minus the probability of incorrectly rejecting a null hypothesis which is true. Many statistical tests are applied at the 95% significance level, meaning that the probability of a false positive is only 5% (1 in 20) when the null hypothesis is true.

Statistic - The calculated value of a population parameter from a particular set of observations. This is calculated according to the appropriate rule, given the spatial and temporal limits of the observations to be included. Since it is calculated from real-world measurements, it will incorporate some error due to both sampling and measurement. Thus, a statistic is an estimate of a parameter.

Two uses can be made of a given statistic: it can be computed to estimate the value of a population parameter (parameter estimation), or it can be tested to see if it meets some pre-determined criteria (hypothesis testing). In practice, parameter estimation and hypothesis testing are slightly different ways of drawing conclusions from the same procedures, depending on the purpose. In this document, the main emphasis is on hypothesis tests, but the same principles can be applied to parameter estimation problems.

STEL - Short-term exposure limit.

Type I Error or False Positive - The type of error which is made whenever a true null hypothesis is rejected. The probability of this type of error is called the significance level of a test.

Type II Error or False Negative - The type of error which is made whenever a false null hypothesis is not rejected. One minus the probability of a Type II error is power.

Variable - A characteristic of an object or system under study. Examples are length, weight, temperature, concentration of a certain chemical, etc. A variable can be thought of as the property that is quantified (to a greater or lesser extent) by a measurement system.

Variables are often classified as independent (or experimental) and dependent (or response) variables. A dependent variable is one which is of interest as a direct or indirect indicator of some process or effect. An independent variable is one which is manipulated by the experimenter in a controlled situation or one that is presumed to have an effect on the dependent variable. The point of a research or monitoring program is to obtain some measure of this effect on the dependent variable.

VOC - Volatile organic compound.

TECHNICAL REPORT DATA
(Please read Instructions on the reverse before completing)

1. REPORT NO.	2.	3. RECIPIENT'S ACCESSION NO.
4. TITLE AND SUBTITLE Guidance for Applying the Data Quality Objective Process for Ambient Air Monitoring Around Superfund Sites (Stages I and II)		5. REPORT DATE August 1989
		6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S)		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Research Triangle Institute Research Triangle Park, N.C. 27709		10. PROGRAM ELEMENT NO.
		11. CONTRACT/GRANT NO. 68-02-4550
12. SPONSORING AGENCY NAME AND ADDRESS Office of Air Quality Planning and Standards US EPA RTP, N.C. 27711		13. TYPE OF REPORT AND PERIOD COVERED
		14. SPONSORING AGENCY CODE

15. SUPPLEMENTARY NOTES

16. ABSTRACT

This document provides direction and assistance on how the data quality objectives process is used to design an ambient air monitoirng system around a Superfund site. Examples are given for Stage I (preliminary decision) and Stage II (refinement of the decision and requirements) for monitoring during remedial action at a hypothetical Superfund site.

17. KEY WORDS AND DOCUMENT ANALYSIS

a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group
Ambient Monitoring Superfund Air Pathways Data Quality Objectives DQO Process Superfund Monitoring	Superfund Air Pathways	

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES 52
	20. SECURITY CLASS *(This page)*	22 PRICE