Chemical

SAFETY AND SECURITY TRAINING

Quality Assurance and Quality Control of Environmental Measurements

Bangkok, Thailand

January 2010

SAND No. 2010-0125P
Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

Sandia National Laboratories

The late 20th Century Became a Time to Focus on Man Made Pollution

Air Pollution due to Increased Vehicle Traffic and Congestion Along with Open Burning

- Combustion
  - SO_2, NO_2 – Air in combustion, heteroatoms in fossil fuels
  - CO – Incomplete combustion
  - O_3- Photochemical smog
  - Pb –fuel additive, Hg- Heteroatom
- Air toxics – incomplete combustion
  - Dioxin
  - Furans
- Suspended particulate matter
  - Dust from unpaved roads
  - Soot – PAH adsorbed

Incineration is not the Same as Open Burning

	Open Burn (µg/kg)	Municipal Waste Incinerator (µg/kg)
PCDDs	38	0.002
PCDFs	6	0.002
Chlorobenzenes	424150	1.2
PAHs	66035	17
VOCs	4277500	1.2

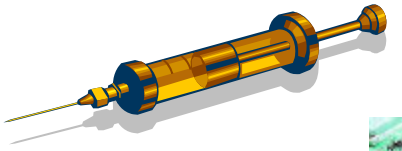
Source: EPA/600/SR-97/134 March 1998

Solid Waste can Directly Impact Human Health

- Decaying animal and plant matter
- Heavy metals
- Solvents – Gasoline, diesel, chlorinated
- Leachates – acid waste, heavy metals
- Hazardous waste – metals, paints, solvents, pesticides
- Leaking fuel tanks – Gasoline, diesel
- Hydrocarbon and oil spills

Uncontrolled Dumping of Medical Waste is a Serious Problem

- Infectious waste
  - Blood,
  - Body parts
  - Body fluids
- Radioactive waste - sources
- Sharps – scalpels, needles, catheters
- Chemicals
- Pharmaceuticals - Narcotics



Industrial Effluents Include Water, Soil and Air Pollution

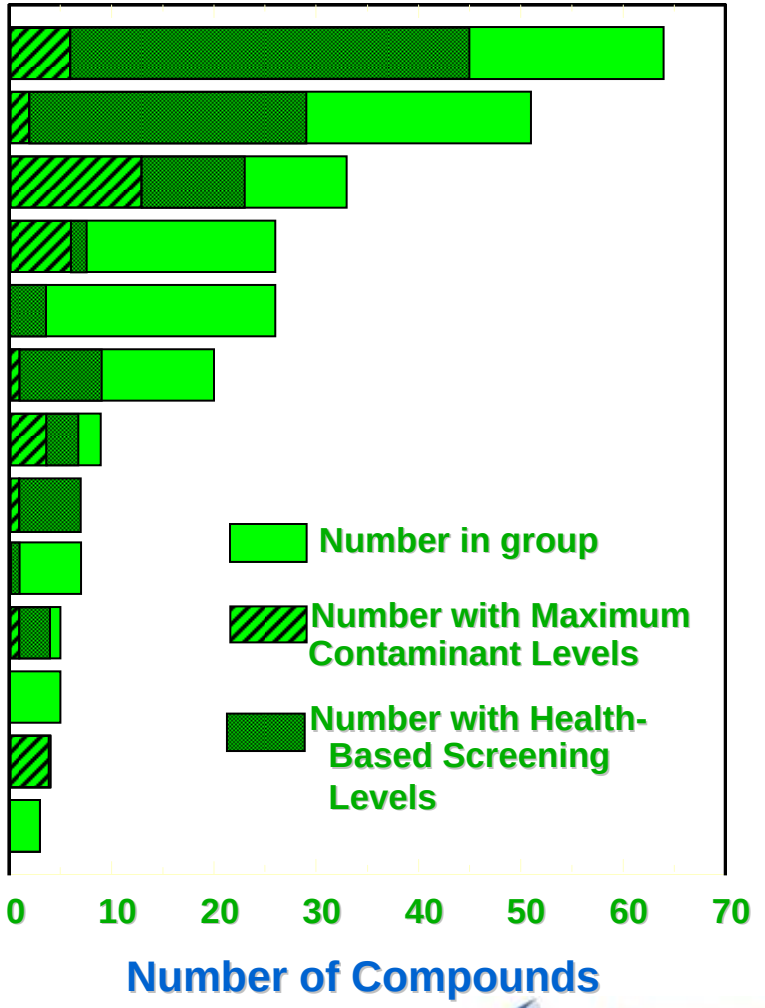
- Tannery- metals, dyestuffs
- Battery manufacture – Pb, acid
- Oil refinery – hydrocarbons, amines
- Chemical processing
- Pesticides – chlorinated, bioaccumulative
- Milling / Mining – Acid mine drainage, sulfide oxidation, mercury used in extraction

Surface and Groundwater Contamination Leads to Health Problems, Water Shortage

- Mining
  - Acid mine drainage
  - Mercury extraction (informal)
- Industrial pollution
  - Dyes and pigments
  - Oil and grease
  - Heavy metals – Hg, Cr, Pb
- Agricultural runoff
  - Pesticides
  - Nutrients – nitrates, phosphates
  - Salinization – Sodium, chloride
- Sewage
  - Pathogens - Enteric
  - Nutrients – Nitrates, phosphates
  - Contaminated animal feed

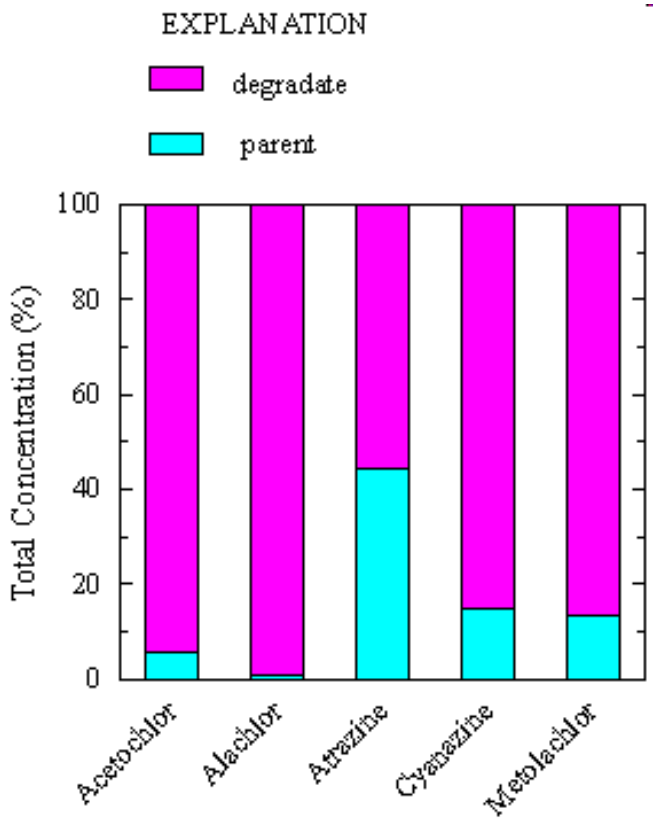
Organic Compounds Monitored

Herbicides and degradates
Insecticides and degradates
Solvents
Gasoline related
Personal care and domestic use
Organic synthesis
Fumigant related
Fungicides
Manufacturing additives
Pavement- and combustion-derived
Plant- or animal-derived biochemicals
Disinfection by-products
Refrigerants and propellants



Source: USGS

Pesticide Residues Depend on Degradation and Prior Use



Common **herbicides** degrade in the environment

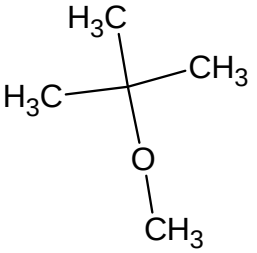
Chlorinated **insecticides** are slow to degrade (aldrin, chlordane, DDT, dieldrin, endrin, heptachlor, mirex)

Other compounds are now used (organophosphorous)

Source: USGS

Petroleum Products are Everywhere - Easily Degraded

- BTEX – Benzene, Toluene, Ethyl Benzene and Xylene
- MTBE – Very water soluble



- Used motor oil
- TPH – total petroleum hydrocarbons
    - Gasoline
    - Diesel
    - Crude oil

Drinking Water, Wastewater Contaminants Directly Affect Public Health

Pathogens
- **Bacteria – Enteric, fecal**
- **Protists – Cysts and spores**
- **Virus - Enteric**

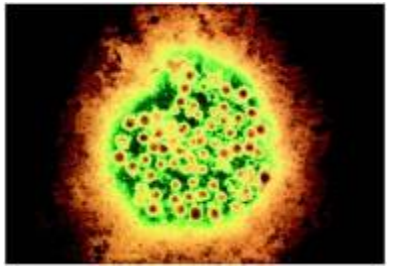
Metals
- **Copper**
- **Lead**
- **Arsenic**

Disinfection byproducts
- **Trihalomethanes - $CHCl_3$,CH_2Cl_2, CH_2ClBr**
- **Haloacetic acid – CH_2ClCO_2H**

Pesticides

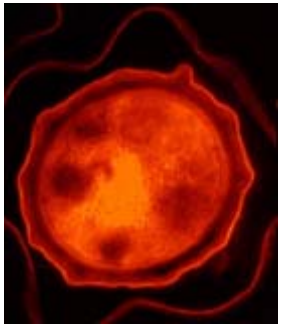
Microbiological Contaminants are Most Important in Public Health

Cholera



Hepatiitis A

Amoeba

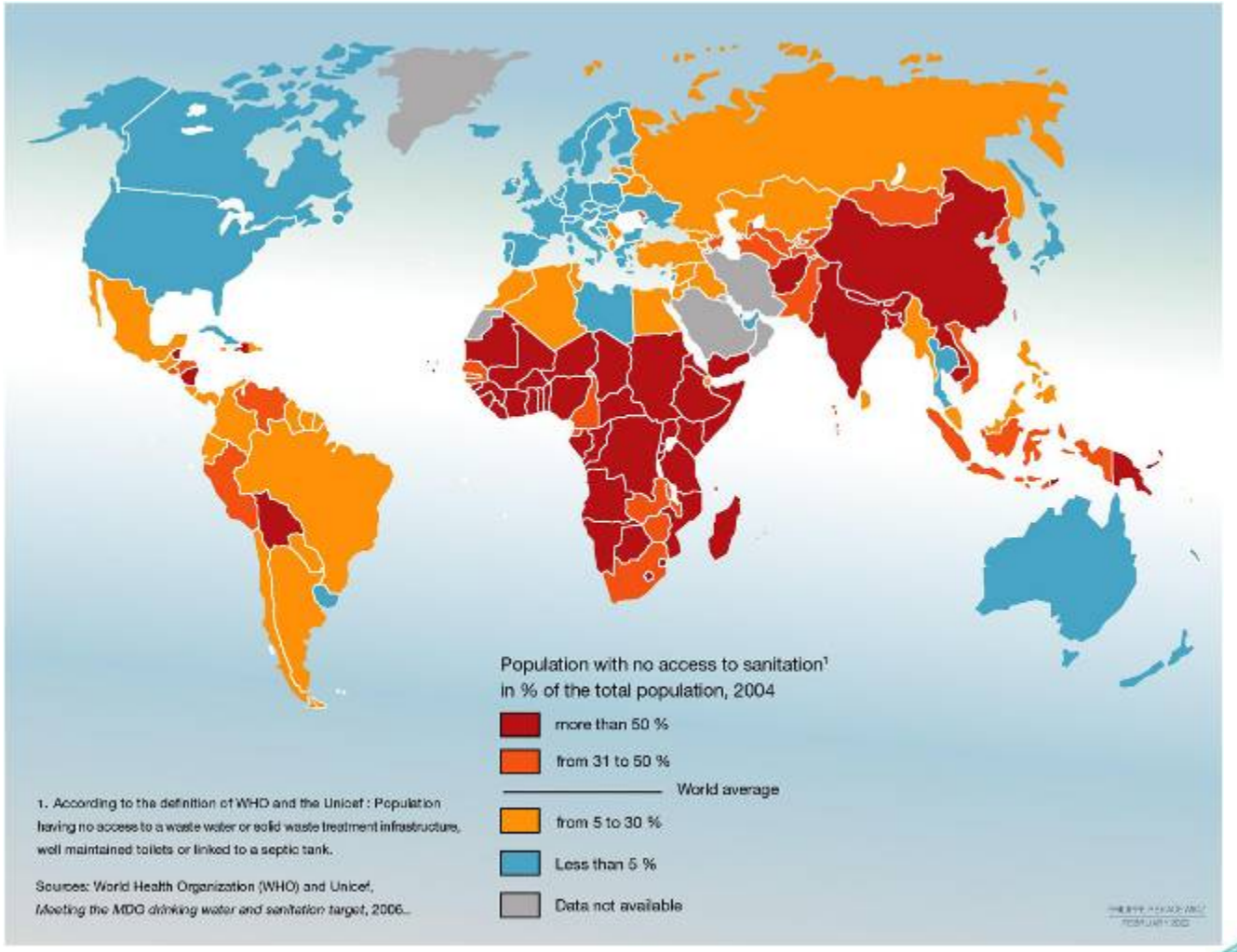


Anthrax spore

Cryptosporidium

Giardia

Lack of Water Sanitation is a the Leading Cause of Early Mortality



Source: WHO

Characteristics of Environmental Analyses

What you need to know and why, so that the data you obtain will be useful and cost effective.

Characteristics of Environmental Samples

- **Three major characteristics that affect environmental analytical data:**
  1. **Analyte concentration levels**
     - **Typically very low (e.g. - a part per billion or less)**
     - **How small is a part per <u>billion</u>?**

  2. **Multiple pollutants may be involved**
     - **Each pollutant may have very different properties, thus requiring different sampling and analysis procedures.**
  3. **Media and matrix diversity**
     - **Media include water, air, solids, wastes**

Analyte Concentration Levels

- **Trace levels of chemical pollutants have deleterious ecological & human health effects, so methods pushed to low detection levels.**
  - Regulations followed to lower & lower concentrations.
  - Many analytes routinely analyzed at part per billion levels.
  - Some, (e.g. chlorinated dioxins and dibenzofurans) routinely analyzed at part per trillion levels (i.e., 10^{-12} g/kg).
  - Some newly targeted endocrine disrupting chemicals (EDCs) may need to be analyzed at hormonal concentration levels which are typically at sub part per trillion levels (perhaps as low as 10^{-15} g/kg)

Multiple Chemicals

- **Sampling may involve multiple chemicals.**
  - **Only a few to 25, 50, 100 or more.**
    - **Many methods developed for sampling as well as for analysis of pollutants.**
    - **e.g. methods for single or multiple elements, volatile organic chemicals (VOCs), semivolatile organic chemicals (SVOCs), chlorinated pesticides, etc. are all different because the properties of the pollutants are very different from each other.**

Media and Matrix Diversity

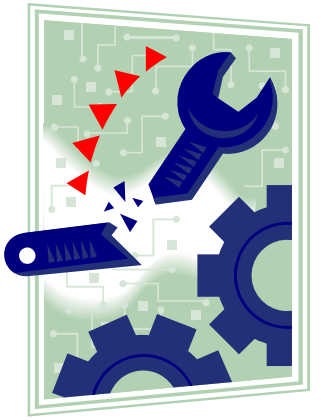
- **Media - water, solids, air, chemical waste.**
- **Additional diversity exists even among media.**
- **Diverse types of media are called matrices.**
  - **Water samples may consist of various matrices such as drinking water, ground water, industrial waste water, rain water, and surface water.**
  - **Surface water can be sea water or fresh water.**
  - **Solids may be soil, dust, wet or dry sediments.**
  - **Air may be indoor air or outdoor (ambient) air.**
  - **Waste may be solids and liquids of all types (chem., plastic, bottles)**

Remember These Questions:

- What are the potential exposure pathways?
- Who or what could be affected by the contamination or by efforts to clean it up?
- What happens if there is “bad” contamination?
  - What happens…if it poses a community threat? If it comes from an adjacent site? If it is migrating off site? If it is from a naturally occurring source?
- Are infrastructures potentially contaminated?
  - Sewers, roads, etc.

Common Problems Encountered

- **Cost overruns**
  - **May occur in sampling & analysis component**
  - **e.g. Inadequate data led to poor decisions.**



- **Failure to meet regulatory and/or community requirements**
  - **May occur due to inadequate site cleanup or inadequate method selection during Site Investigation activities.**
    - **Samples taken in wrong places and/or not enough samples taken,**
    - **Methods used did not have good enough sensitivity, selectivity, accuracy, or precision.**

A Common Problem With Field Analysis Projects

- **Projects with cost & time overruns and results that do not adequately support the decisions for which the data was generated often require re-sampling and re-analysis.**

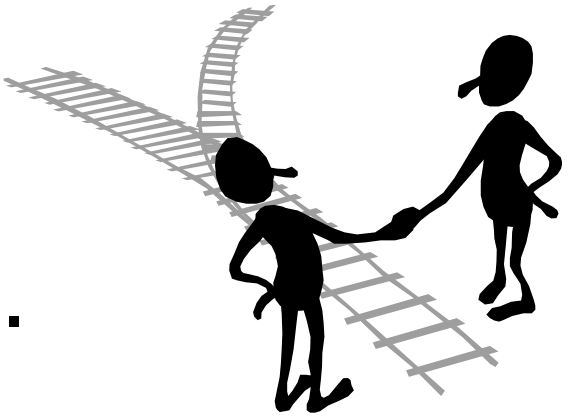
- **Solution: implement systematic planning with confidence levels for key measurement criteria.**
  - **Confidence levels - statistical calculations of probability of data being correct.**
    - **Serve as a way to quantitatively measure quality of data.**
    - **Is the probability 95%, 90%, 80%, 50% or what?**
    - **Unless you can calculate a confidence level you do not know how probable the data will be correct.**

How Can Data Quality Be Guaranteed?

- **When all the key criteria of interest are defined and agreed to,**

- **And when the sampling and analysis plan is developed to meet these criteria at the agreed upon confidence levels,**

- **And <u>IF</u> the plan is followed, the data quality can be guaranteed.**



What May Cause Failure?

- Workers or laboratories not following the plan
  - Wrong types of samples taken or taken at wrong location
  - Fewer than agreed upon numbers of samples taken
  - Samples not collected and/or preserved correctly
  - Samples not analyzed according to plan
  - Analytical interferences in samples
  - Wrong kinds of QA/QC samples taken
  - Insufficient numbers of QA/QC samples taken

What are the Consequences of Failure?

Consequences vary with the type of failure,

BUT

If you incorporate confidence levels into systematic planning, you can quantitatively measure the consequences of failure and decide if the data is still useful. If you can't use the data you will be able to determine how to fix the problem.

Factors That Affect Data Quality

The quality of environmental analytical data is affected by four major factors:

1. **Variability in data - affected by representativeness of the samples. (how well does 1 liter of water represent a lake, a river, a drinking water system?).**

2. **Sampling factors, i.e., difficulty obtaining them. (How many samples are collected, how different they are?)**

3. **Analytical factors, i.e., multiple methods with different detection levels, accuracy, precision and selectivity for a chemical.**

4. **Data handling factors, i.e., how data are mathematically treated can distort the evidence.**

Variability in Data

- There are multiple sources for error and variability in sampling and analysis.

- Sampling components usually form largest source.
  - Including the distribution of pollutants in the sample itself.
  - In spite of the this, most QA/QC focuses on analytical and data handling components.

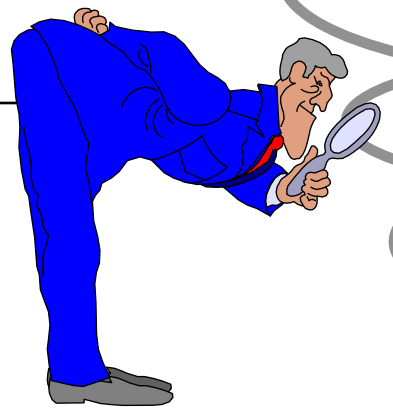
There are Multiple Sources of Variability Which Affect Data Quality

Pre-Lab

Pollutant Distribution

Field Sampling

Sample Handling



Sample Prep.

Analysis

Data Handling

After Lab

Where are most QA/QC efforts focused?

Samples Must Represent a Site

- **The first consideration for evaluating data is whether the samples adequately represent the site.**
  - **Remember, the purpose of environmental sampling is to obtain a *small* but *informative* portion of the media being investigated.**
  - **Samples are thus representative of the site and conclusions involving data as evidence are based on them (a very small sample represents a very large site – this is very difficult to do if you think about it).**

What happens if samples do not represent the site?

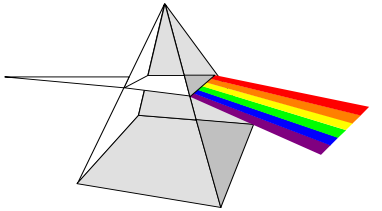
“Representative” May Mean Different Things

- **Samples must be representative of the site, but not necessarily “representative” samples in the statistical sense.**
  - ***Judgmental sampling* provides data from a site that is representative of an area.**
    - Not representative of the whole site
  - **Often used to obtain “worst case” scenarios involving pollution.**
    - Examples include sampling a visibly contaminated spot on the ground or around a structure used to store or handle chemicals, biological agents, or other hazardous materials.

Pollutant Distribution In Samples

- **Pollutant distribution in samples is usually the largest contributor to analytical variability.**

- **Pollutants are not distributed evenly in the environment.**
  - **Even distribution is called "homogeneity."**
  - **Non-homogeneity is the expected case.**
  - **Sometimes efforts are made to make the samples homogeneous after collection. This is useful for testing the quality of the analytical method (but it does not change the non-homogeneity of the site that is sampled).**



Un-even Pollutant Distribution Exists Everywhere

- In soils and solids non-homogeneity (heterogeneity) is expected.
- However, in water it is also common.
  - Stratification usually occurs along shorelines, at different depths, and at different times.
- Even in air non-homogeneity is very common.
  - Stratification in plumes carried by wind is typical.
  - Temperature inversions also cause stratification.

Pollutants Usually Don't Stay in One Place

- Another complicating factor is the fact that pollutants migrate (move) in the environment.

- Migration occurs rapidly in air and with more variability in water.
    - May be rapid in rivers and moving streams, slower in lakes and ponds, and slower still in groundwater sources.
    - Migration of pollutants in land sites results from movement through groundwater or by movement down through various layers of soil from rain.

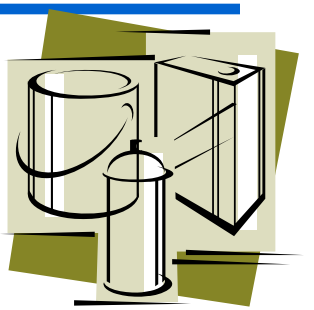
Sampling Factors That Affect Data Quality

- **The second consideration for evaluating environmental data is sample collection.**

- ***Small*** **portion of media from a site is placed into a container for analysis or transportation to a lab.**
  - **Improper containers result in contamination & analyte loss.**

- **Usually samples are preserved (often by cooling with ice) and then submitted to field instruments or a lab.**
  - **Improper preservation or storage of samples result in degradation & analyte loss.**
    - Volatile pollutants (e.g., TCE, chloroform, etc.) have many requirements such as storage time, place, and temperature.

Different Types of Pollutants Have Specialized Sampling Requirements



- **<u>Biological</u> - *E. coli* and other bacteria.**
- **<u>Inorganic</u> - mercury, lead, and other metals.**
- **<u>"Characteristic" or "Water Quality"</u> - BOD, COD, nitrogen, phosphorus, etc.**
- **<u>Organic</u> - all contain carbon atoms**
  - **Volatile Organic Analytes (VOAs) or Volatile Organic Contaminants (VOCs) - TCE, chloroform, etc.**
  - **Semivolatile Organic Contaminants (SVOCs) - PCBs, most pesticides and industrial chemicals.**

Sample Collection for Chemical Analysis Depends on Types of Pollutants Collected

- Sample containers may cause degradation or interferences with target analytes.
  - Plastic is used for water samples with metals.
  - Glass is used for water samples with organics.
  - VOCs require vials with Teflon septa in the caps.
  - Light-sensitive organics require amber glass.
  - Similar containers used for solid samples as for water.
  - Air samples are collected in passivated steel canisters, Tedlar bags, and on various solid sorbents.

Preservation & Storage Factors May Affect Analytical Results

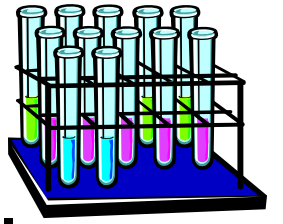
- Samples are usually preserved and transported to a lab for storage unless analyzed in the field (field methods or mobile lab).

- Large diversity of preservation & storage conditions for water samples.
  - Dependent on the specific chemicals.
  - Diversity in other media: soils, sediments, wastes, and air.

- Questions to ask:
  - Were the samples preserved (and how)?
  - Were the samples transported and stored in a way that protected them from degradation or contamination?

Detection Levels Affect Methods Used and Amounts of Samples Needed

- **Desired detection levels may affect amount of sample needed for analysis.**
  - Example: larger amounts of low level samples may be prepared for concentration in a lab using standard laboratory methods.



- **Detection levels may also affect the way samples may be collected.**
  - Example: multiple low level samples composited together could result in false negatives (target analytes are present but they are missed by the analysis due to dilution from compositing multiple samples).

Composited Samples

- **A composite sample is made up of smaller portions of samples taken over either time or space to provide a single sample for analysis.**

Advantages:

- Less expensive analysis
- Average results over time or space.

Disadvantages:

- Higher detection levels
- Higher probability for false negatives
- Can not determine variability with time or space.

Analytical Factors That Affect Data Quality

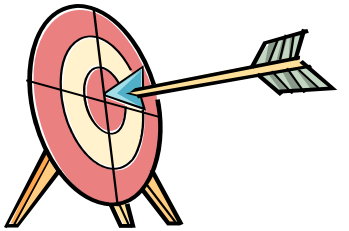
- **Method Sensitivity**
  - **If a method is selected that is not very sensitive to the target chemicals then they may not be <u>detected</u> even though they are present.**
    - **"False Negative" meaans present but nort detected**

    In addition, the concentration level at which they can be <u>measured quantitatively</u> will be higher than if a more sensitive method was used.

  - **Most regulatory decisions are based on reports that show presence of pollutants above a regulated concentration.**

Other Analytical Factors That Affect Data Quality

- **Selectivity of the method used**
  - **Insufficient selectivity → false positive results**



- **Accuracy of the method used**
  - **Inaccuracy may result in data not being accepted**
- **Precision of the method used**
  - **Imprecision may result in data not being accepted**
- **Details will be covered in Method Selection Slides**

Data Handling Factors That Affect Data Quality

- Uncertainty levels that are higher than Data Quality Objectives (DQOs – discussed later) may be caused by:
  - Insufficient number of samples
  - Detection levels that are too high
  - Accuracy and precision values that are too low due to methods selected with insufficient performance.

- Improper reporting by the laboratory
  - Failure to properly "flag" data that do not meet requirements of the analytical system.

- Audit failures by a laboratory

Summary

- **The samples must be *representative* of the site**
  - **Pollutants are unevenly distributed and migrate**

- **Obtaining *reliable* environmental data is difficult**
  - **Sampling factors usually outweigh analytical factors**
  - **Most of the QA/QC is focused on analytical factors**
  - **It's worse than looking for a needle in a haystack**
  - **Many non-target chemicals are in samples**
  - **Usually multiple media must be considered**
  - **There are many procedures where errors may occur**

Why QA/QC is so Important

Lawrence H. Keith, Ph.D.
Environmental & Chemical Safety Educational Institute
Monroe, Georgia

Points Stressed

- **What you see may not be what you have.**
  - **When a number might be a different number**
  - **Could 18 ug/L and 23 ug/L be the same number? (Yes)**

- **There is more to an analysis than the lab data.**
  - **It's the sampling! (When sampling is bad it doesn't matter how good the lab data is – you still get the wrong answer!)**

- **How very extremely tiny a part per billion is.**

- **Fast, Good, Cheap - select any two.**

- **Plan twice, measure once.**
  (Good planning is important!)

Commonly Encountered Issues and Decisions

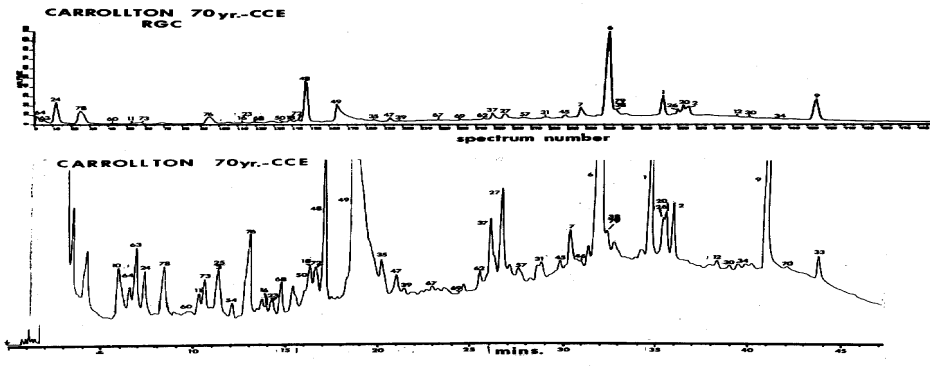
Important Concepts That Affect All Environmental Sampling and Analysis Projects

Commonly Encountered Issues

- **All environmental analytical measurements have errors in their reported values.**
  - **"*Error*" is another word for "*variability*" and is used by statisticians.**
  - **It is the difference between measurement result and the true (but unknown) value of what is being measured.**
  - **One of the most important things to know about analytical data (is a value + or – 10%, 20%, or 50%?)**
  - **Error is discovered and measured by analyzing QC samples (are many different kinds of QC samples).**

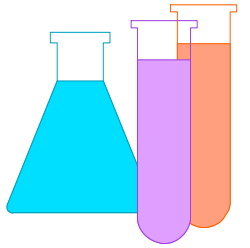
Commonly Encountered Issues

- **Pollutants are at very low concentrations (ppm/ppb)**
- **Pollutants migrate - especially in air and water and this contributes to analytical data variability.**
  - **The number you get may not be the number you have!**
- **Most pollutants are not <u>directly</u> measured (they are signals drawn on paper or entered in a computer).**



New Orleans Drinking Water Study (Circa 1974)
GC/MS Ion Current Chromatogram with Identified Compounds Noted (Upper Trace)
GC-FID Chromatogram with same Identified Compounds Noted (Lower Trace)

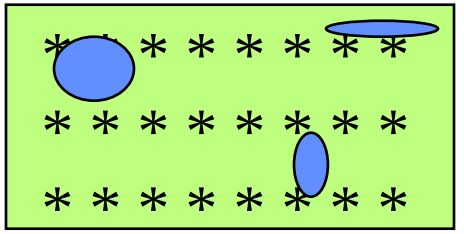
Commonly Encountered Issues

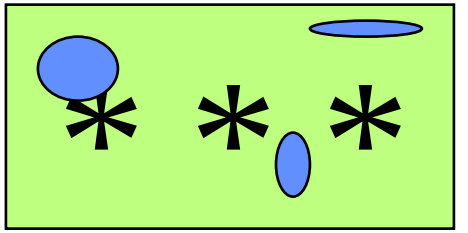


- **Pollutant distributions are *not* homogeneous**
  - **e.g. pollutant concentrations in soil, water, and air may vary in the field by time and/or physical place.**
- **Samples you collect represent the field site**
  - **A liter may represent thousands or millions of liters of water at the field site (part of a lake, part of a river, etc.)**
  - **How difficult is it for that tiny portion to be representative?**
- **And QC samples represent the field samples**
  - **What happens if your field samples do *not* represent the site and/or if the QC samples do *not* represent the field samples**
  - **Have you wasted time and money? (Probably)**
  - **What could the consequences be? (Incorrect Decisions)**

Lesser Known Commonly Encountered Issues

- **Definitive methods do not guarantee good data**
- **High quality data is often less important than data with high information content**
  - **Remember the inhomogeneity of field sites**





Cheaper Analyses With Lower Quality Data = More Samples and Higher Information Content With the Same Budget

More Expensive Analyses With Higher Quality Data = Fewer Samples and Lower Information Content With the Same Budget

Lesser Known Commonly Encountered Issues

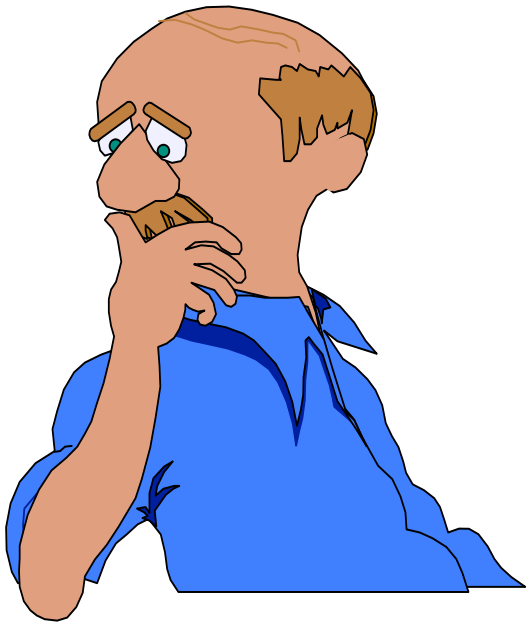
- The closer you measure to the detection level of a method the more QC samples will be needed.
  - More samples = *$ HIGHER COSTS $*
- Chemical detection and chemical identification are separate issues involving different requirements (detectors & QC samples).
- Chemical detection and quantification are separate issues with different requirements (quantification is 10x detection)
- Method selection affects reliability and cost (usually there are many methods to choose from).
  - Selectivity affects pollutant identification
  - Sensitivity affects detection and quantification
  - Accuracy and precision affect data quality

  Each of these are characteristics of every method!

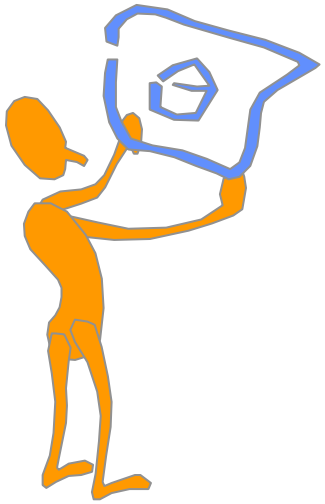
Each Issue Will be Addressed

- We'll learn about them in the context of field sampling and analysis.

- We'll learn about the new way of working smarter and cheaper instead of harder with the old concepts and ways.



What You Need to Know Before Implementing Work



Before You Begin…

- **How will the site be used in the future?**
  - **This will affect most field projects, for example:**
    - **Site investigations & cleanup requirements**
    - **Cost of sampling and analysis**
    - **Cost of cleanup options if a site is contaminated.**
  - **Defining & understanding goals and decisions is critical to identify appropriate technologies for site investigation and cleanup.**
    - Goals affect collecting data & decision-making.
    - For example, will the site be used for industry or for a living area?

Understanding Objectives

- **What if you don't know the end use?**
  - Try to identify general type of development (industrial, residential, mixed use).
  - If no idea is available, then conservative assumptions and actions should be made at each stage.
    - **This is probably the most costly option.**

- **Understand Pakistan's applicable environmental laws & regulations.**
  - Critical for selecting technologies for cleaning up sites.
  - Pakistan's EPA authority may drive site requirements.
    - **Assessment, cleanup options, waste management, etc.**

Site Assessment

- **Objective: Evaluate potential for contamination**
  - **Collect and review existing information.**
    Search historical records
    - Need past and current uses of the site.
    - Gather past and current environmental conditions.
  - **Critical step - work hinges on whether potential environmental concerns are identified.**
    - End use of the site determines types and quantity of data that must be collected as well as quality of data.
    - **Data Quality Objectives (DQOs)** serve as basis on which decisions on which technologies will be used.

Data Quality Objectives

- ***DQOs*** **are qualitative and quantitative statements used to ensure that data of known and appropriate quality are obtained.**
  - DQO process is a series of planning steps.
    - Typically conducted during site assessment and investigation.
    - Designed to ensure that appropriate data (and also that sufficient numbers of data) are used in decision-making.
  - Necessary because many different technologies available with varying degrees of quality, cost, and difficulty of use.
  - It lets everyone know how “good” the data needs to be.
  - More on this later…

Site Assessment Uses

- Site assessment can provide preliminary indication of what types of cleanup technologies might be available.

- Site assessments are essential to assess and address needs and concerns of the community.
  - For example, development of social and economic profiles and identification of acceptable environmental risk by the community.

What Do You Need to Know From Site Assessments?

- **Is site located in an area for redevelopment?**
- **Is site in an industrial area?**
  - If so, will it remain such or be changed to residential use?
  - Will community members be exposed to soil or sediment?
- **What data is needed to support goals / concerns?**
  - What level of quality or uncertainty is necessary? (DQOs)
- **What records exist about use / contamination?**
  - Complaints filed, environmental audits?
  - What is needed to know extent or absence of pollutants?

Actions From a Site Assessment

- **Document what is to be accomplished at a site.**
  - Redevelopment, cleanup, etc.

- **Document what decisions are needed for goals.**
  - Which contaminants should be tested for, what are acceptable remaining concentration limits, and what is the necessary degree of certainty (95%, 90%, 80%, etc.)?

- **Identify data or information needed to make these decisions or answer the questions above.**
  - Determine what level of data quality is needed for each pollutant of concern (accuracy, precision, confidence level).

Actions From a Site Assessment

- **Develop and evaluate initial plans for sampling & analysis to achieve required data and data quality.**
  - Consider technologies available to provide more data, quickly and cost-effectively, at quality level needed.

- **Determine what the data will tell you when you have it:**
  - Can you make the proper decisions based on it?
  - Will all your questions be answered?
  - Will you have the quality needed (not too much or too little)?
  - Can you afford it (both time and money)?

How Do You Find the Answers?

- **Establish a technical team covering all disciplines**
  - If chemical pollutants are involved make sure an environmental chemist is on the team - often not done!
- **Ensure all decision makers are involved in planning.**
  - Regulators, citizens, property owners, technical staff, etc.
- **Identify the project goals and data needed to support them - (write them down).**
- **Determine what contamination, if any, is likely**
  - Search records and examine the site - but no sampling yet.

How Do You Find the Answers

- Conduct record searches such as sewer maps, aerial photos, health documents.
- Review any government records for site use.
- Review any government lists of site problems noted.
- Interview previous owners, employees, neighbors.
- Perform visual exam of land and buildings.

Key Site Assessment Steps if:

- **No apparent contamination**
  - **Discuss results with key officials before proceeding.**

- **Contamination is found that poses significant risk to human health or the environment.**
  - **Determine levels of cleanup needed (for example, based on planned use of the cleaned site). Then proceed to Site Investigation step.**

- **Contamination possibly exists - or it exists but no site investigation has been conducted or more investigation is needed.**
  - **Proceed to Site Investigation step.**

Site Assessment

- **Objective:** Identify the source, nature, and extent of contamination.
  - Detailed study is performed.
  - Results are used for field plans to:
    - Determine goals for cleanup (minimum levels remaining).
    - Determine risks (to humans and the environment) based on the pollutants and how much of them are present.
    - Determine acceptable risks (what amount of residual pollutants left at the site will be acceptable for site use?).
    - Develop effective cleanup plans (how will cleanup be done?).
    - Determine type of data to be collected to support plans.
    - Determine level of uncertainty acceptable in the data.

What Do You Need to Know?

- Which agency will be responsible for oversight?
- Who will review the Site Investigation Plan?
- What technologies are available for data collection?
- Has the technical team explored different methods?
- Will selected technologies provide all data needed and at the desired confidence levels in DQOs?
- Will government or community issues affect the investigation?
  - How should these be addressed?

What Do You Need to Know? (2)

- **What are the potential exposure pathways?**
  - **Air, Water, Food, Skin**
- **Who or what could be affected by the contamination or by efforts to clean it up?**
- **What happens if there is "bad" contamination?**
  - What happens… if it poses a community threat? If it comes from an adjacent site? If it is migrating off site? If it is from a naturally occurring source?
- **Are infrastructures potentially contaminated?**
  - Sewers, roads, railroads, water supply, etc.

How Do You Find the Answers? (1)

- **Identify proper mix of technologies:**
  - **field measurements and fixed lab measurements**
  - **If chemical pollutants are involved make sure an environmental chemist is on the team.**
  - **Ensure lab and field instruments have needed detection levels, specificity, accuracy & precision for their uses.**
    - Some field instruments provide screening level data and some higher level data - a mix plus higher level data from fixed lab methods often provides the most *cost effective* solution.

How Do You Find the Answers (2)

- Determine conditions at the site.
  - Use site assessments to determine identities and locations of pollutants
  - Follow with laboratory sampling & analysis to confirm identities and concentrations of pollutants as necessary

- Assess risk to human health & environment
  - Consider exposure pathways:
    - Inhalation, ingestion, skin contact of soil & dust.
      - Ingestion of water.
      - Inhalation of air.

Characterization Steps

- **If site assessment is needed then characterization of the site will be necessary, and…**
  - **You will need to understand characteristics of environmental analysis and common problems encountered.**
  - **You will need to know what factors affect data quality.**
  - **You will need to evaluate "*dynamic*" investigations using dynamic sampling plans for cost effective solutions.**
  - **You will need to know how method selection for analysis of pollutants will effect the quality of the data used to make important decisions and what questions to ask.**

What You'll Learn Next

- Field samples are not homogeneous.
- Samples are taken to represent a field site.
- And then QC samples represent the field samples.
- Pollutants are at very low concentrations (ppm/ppb).
- Pollutants move - especially in air and water.
- Why all environmental analytical data has variability and what to do about it.
- Most pollutants are not <u>directly</u> identified.
- Definitive methods do not guarantee good data.

What You'll Learn Next

- High quality data (accurate & precise) may be be less important than data with high information content .
- The closer you measure to the detection level of a method the more QC samples will be needed.
- Pollutant identification and quantification are totally separate issues with different requirements.
- Method selection affects reliability and cost.
- How to determine how many samples ($$$) you'll need to meet required confidence levels.
  - 95% or 90% or 80% or 50% confidence levels for example

Systematic Planning

Data Quality Objectives & Method Quality Objectives
or
How To Avoid Nasty Surprises

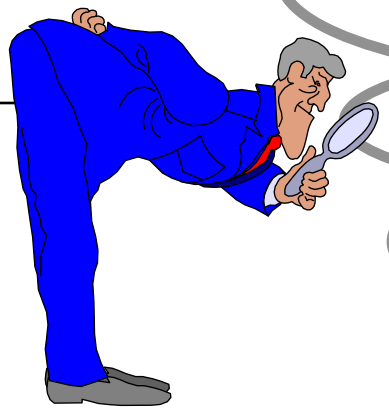
Remember This Slide About Sources of Variability?

Pre-Lab

Pollutant Distribution

Field Sampling

Sample Handling



Sample Prep.

Analysis

Data Handling

After Lab

Where are most QA/QC efforts focused?

The Purpose of Environmental Sampling and Analysis

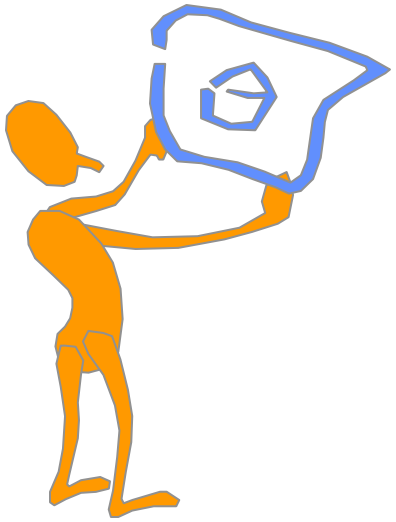
To obtain a small but informative portion of the population (environmental site) being investigated.

You can’t sample the whole environment. Therefore, samples representative of the site are taken and conclusions are made based on these selected samples.

Look at the Whole Picture

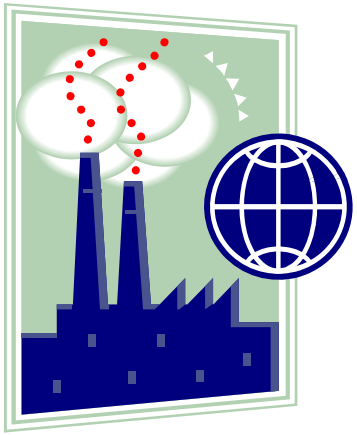
- **Systematic planning should consider both sampling and analytical activities.**

- **Using EPA's Data Quality Objective (DQO) process is the most common form of systematic planning.**



Who Benefits From Using Systematic Planning?

Anyone who generates or uses environmental information to make important decisions can benefit from using the DQO process.



- **Examples include:**
  - Engineers, hydrologists, ecologists, and others who manage environmental projects where decisions are based upon use of information generated from field sampling and analysis activities (which may be subcontracted).

How Could *You* Benefit From Using the DQO Process?

- **Government Administrators –** increase efficiency, save time, and be able to quantitatively measure employee's progress and job performance

- **Industrial Managers & Directors –** decrease production costs and improve products by minimizing error sources and evaluating other ways of working that will increase production

- **Laboratory Managers & Analysts –** improve analytical performance by finding critical steps that can be revised, reducing false positive and false negative identifications, and using QC samples that are specific for each kind of error source as needed when they are needed.

How Can *You* Save Time and Money Using Systematic Planning?

(1) By being able to evaluate the quality of the information that will be obtained from environmental (*or other*) projects before collecting samples and analyzing them, and…

(2) By producing an optimal study (*or production or other work*) design meeting your needs the first time a project is executed – *so it doesn't have to be repeated!*

- **This is important because field activities are typically expensive and often generate data that do not completely meet the project's needs.**

How Else Can *You* Save Time and Money Using Systematic Planning?

- **Achieve cost and time reductions when planning, implementing, and assessing environmental projects.**
  - **Eliminate repetition of all or portions of information collection (e.g., re-sampling and re-analysis – this is very expensive!).**
  - **Conduct only those field and analytical data collections necessary to result in statistically required confidence levels (not too many or too few since both waste time and money).**
  - **Ensure that important factors will not be overlooked during planning (Don't realize you needed other data at the end).**
    - **Guarantee the information being collected will have the required degree of quality for its intended use initiating field activities.**

Data Quality Objectives
and
How to Use Them

How Can You Avoid Nasty Surprises?

Use EPA's Data Quality Objective Process

- **The DQO process is an iterative (*repetitive*) and flexible procedure**
  - **Produces both qualitative and quantitative statements.**
  - **These define data quality criteria and information design requirements.**
  - **Using a systematic seven-step iterative procedure ensures that all the information needed is acquired, with the degree of quality that is required, *the first time* a project is executed.**

U.S. EPA Originated the DQO Process

- **U.S. EPA's Quality Assurance Management Staff developed the DQO process as an important scientific and legally defensible data collection and planning tool.**
  - **Were spending millions of dollars and getting data that often did not meet their needs or where the quality of the data was not known (may have been good or bad but nobody could tell which).**
- **Developed to help people decide what type, quality, and quantity of data are needed for environmental decision making.**
  - **Iterative and flexible**
  - **Produces both qualitative and quantitative statements that define data quality criteria.**

DQOs and MQOs

Data Quality Objectives (DQOs)

- **Statements that define the confidence required in conclusions from data.**
  - **A convenient format to state objectives that are both comprehensive and measurable.**
  - **Specify the amount of certainty required (% probability).**
  - **Should be part of every sampling & analysis plan.**

Measurement Quality Objectives (MQOs)

- **Specific values used to measure the quality of analytical data.**
  - **Examples include DLs in ug/kg; SD ug/mL; % RSD; % Recovery.**

Common Problems With Current Approaches to Environmental Analysis

- **Even though environmental sampling is recognized as the weakest link in the chain of events involving *planning* ➔ *sampling* ➔ *analysis* ➔ *data evaluation*, we still focus most strongly on the latter steps.**

- **Even though we know some methods are inadequate, people continue to use them in a *prescriptive* manner.**
  - **Even when some methods are "good" they may not perform well in all matrices (for example, water methods used with soil or plant/animal tissues).**

The DQO Process

- **There are seven (7) steps in the DQO process.**
- **The final step is optimizing QA program design:**
  - **Determines how many samples are required to meet an objective and which type of sampling plan is required (for example, simple random, stratified random, judgmental, etc.) .**
  - **Iteration (*repeating*) of this step is usually necessary to optimize the number of samples required to meet DQOs and meet budget, time and other study requirements.**

Seven Stages of DQO Planning

1. State the Problem to be Resolved
2. Identify the Decision to be Made
3. Identify All the Inputs to the Decision
4. Narrow the Boundaries of the Study
5. Develop a Decision Rule
6. Develop Uncertainty Constraints
7. Optimize the Design for Obtaining Data

What are Some Basic DQOs?

- **What percentage of false positives and false negatives are you willing to accept?**
  - **These are *incorrect conclusions* either about the identity of the analyte or its concentration**
  - **Most analytical factors considered will affect the amount of false positives & false negatives.**
    - **Analyte Recovery**
    - **Detection Levels**
    - **Standards and Calibration Curves**
    - **Method Specificity**
    - **Interferences**

DQO Components

- DQOs are developed using a process that encourages the sequential consideration of relevant issues. <u>You decide on the following</u>:
  - The precise problem(s) to be resolved
  - Decisions needed to resolve the problem(s)
  - Inputs needed to make the decision(s)
  - The boundaries of the project
  - The decision rules (IF this THEN that)
  - The uncertainty constraints (% Probability).

- Examples of these components are next

Seven Steps to Achieve Successful

Project Results

and

How to Use Confidence Levels With Systematic Planning

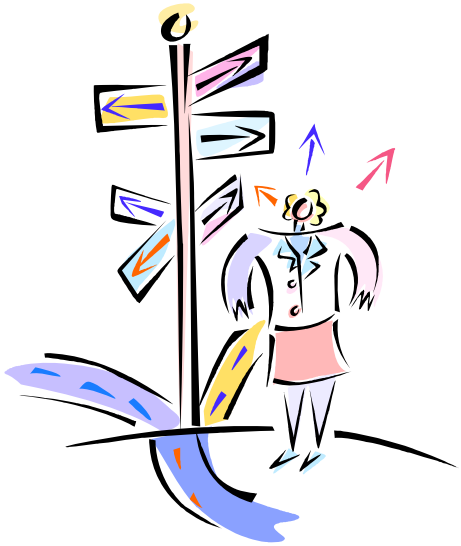
(Step Number 6)

Step 1 - State the Problem to be Resolved

- **Examples:**
  - **To determine if specific pollutants are present at a field site above desired concentration limits.**
    - **The site may involve a land, a pond or lake, drinking water source, monitoring wells, a stack plume, etc.**
  - **To track changes in concentration of specific pollutants at a site over time (e.g. groundwater plume).**
    - **This is more typical with drinking water sources.**
  - **To determine if groundwater needs treatment in order to meet regulatory or safety requirements.**

Step 2. Identify the Decisions

- List each of the decisions that will be made.
  - Decisions involving use of the field site
  - Decisions involving cleanup, if needed.
  - Decisions involving regulatory mandates.
  - Decisions involving community leaders.



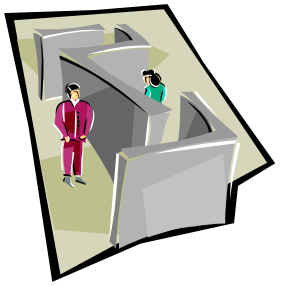
Step 3 - Identify All the Inputs to the Decisions

- **This is often a difficult step – *details* are needed.**
  - **What data is needed to make decisions?**
    - **Physical characteristics (pH, color, turbidity, temp.)? Meteorological (weather) measurements? Specific chemical identifications? Chemical concentrations also? What detection levels? What sample matrices (ground water, industrial wastewater, soil, sediment, air, plants)?**
    - **How will the necessary data be obtained?**
      Site assessments, field analysis, lab analysis, other?
    - **Where are all the places that samples will need to be collected? What sample containers will be needed? What preservation will be needed?**
    - **What methods will be used? Are the detection levels needed achievable? Is the method verified for the matrices? Is it specific enough to identify the pollutants with good confidence (no false positive identifications).**

Step 4. Narrow the Boundaries of the Study

- The two primary boundaries are time and space

  - Time: How long will the study take (as initially planned)?



  - Space: What are the boundaries of the study (as initially planned)?

  - But there is also a third one - Money: Although this is not usually considered a boundary it really is. Most projects have budgets and require funds.

Step 5. Develop Decision Rules

- It is important to develop as many as possible so decisions will be understood by all.
- Decision Rules are IF/THEN statements:
  - IF the study takes more time than planned, THEN …
  - IF expected pollutants exceed regulations, THEN …
  - IF unexpected pollutants are found, THEN …
  - IF pollutants are found migrating off site, THEN …
  - IF the methods selected don't meet requirements for sensitivity, selectivity, accuracy, or precision, THEN …
  - IF stakeholders want more data than planned, THEN..

Step 6 - Develop Uncertainty Constraints

- **Uncertainty constraints are limits on the amount of error in the data that is tolerable.**
  - **They are used to establish performance goals for design of the data collection activity.**
  - **These limits should be based on the consequences of making a wrong decision.**
    - **Consequences causing human health risk should have smaller errors that will be tolerated.**

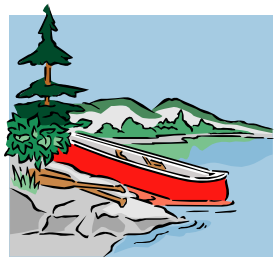
Remember one size does NOT fit all.

Step 6 - (Continued) Develop Uncertainty Constraints

- **Environmental data <u>always</u> has some degree of error associated.**
  - **Decision makers often do not realize this fact.**
  - **Environmental data only provides estimates of the "true" situation.**
    - **For example, you never analyze:**
      - **All the water in a lake or pond**
      - **All the air in a specific location**
      - **All the drinking water in a city**
      - **All the food processed at a site**

Step 6 - (Continued) Develop Uncertainty Constraints

- **Thus, you collect a limited number of samples to analyze and use the data from them to *represent* the rest of the environmental site (land site, river, lake, city air, all the other items of food, etc.).**



- **You use the <u>representative</u> data generated to estimate the probable chemical pollution in the rest of the environmental site.**

Step 6 - (Continued) Develop Uncertainty Constraints

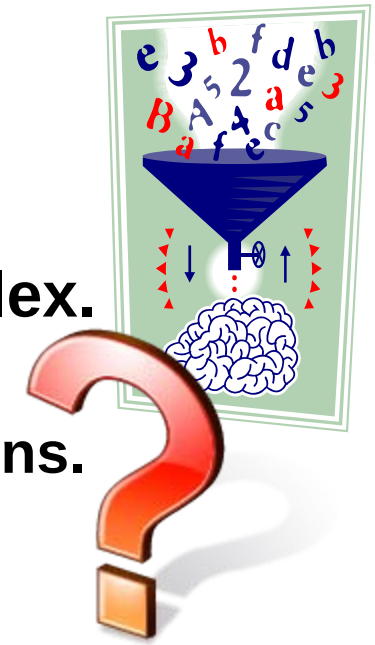
- **There are many sources that contribute to the uncertainty (variability) in environmental data:**
  - **Uneven distribution (*heterogeneity*) of chemical pollutants in the environmental matrix being sampled,**
  - **Error from sampling design,**
    - **Error inherent in using a portion to represent the whole.**
  - **Error from sample collection and preservation,**
  - **Error from analytical measurements, and**
  - **Error from data handling**

Step 6 - (Continued) Develop Uncertainty Constraints

- Uncertainty is measured several ways.
- These form the basis of deciding whether or not your DQOs and/or MQOs have been met.
  - Rates (%) of false positive and false negative conclusions in the analytical data,
  - Precision (closeness of values from repeat analyses - expressed as standard deviation),
  - Bias - influences accuracy - (closeness of measured values to “true” or “assigned” values).

Step 6 - (Continued) Develop Uncertainty Constraints

- **Problem – many environmental projects don't have information available on the data uncertainty.**



- **Because the statistical equations are complex.**
  - **People usually don't know what inputs are required or how to use the statistical equations.**

- ***Result* – much environmental data is of unknown quality.**
  - **Good, bad, or mediocre (for the project DQOs)?**

Step 6 - (Continued) Develop Uncertainty Constraints

- **How do you find out how "good" data is?**
  - **Remember, "good" is relative to field site required quality.**
- **One way - *DQO-PRO*, an easy-to-use statistical calculation program provided free from the American Society (ACS).**
- **Written for the US Government so it is**
  - **It can be copied and distributed without any limitations.**
  - **Dr. Keith is the primary author.**

Step 7 - Optimize the Design for Obtaining Data

- A satisfactory design for obtaining data to meet specific objectives is rarely achieved the first time.
  - Iteration is usually needed.
  - If not with Steps 1 - 5, then usually Step 6.
  - Affects the statistical confidence levels of the data and confidence may have to be decreased to meet budget limitations.

How Confidence Levels Affect Project Costs

Numbers of Samples = $$$

- The greater the confidence levels required the greater the number of samples needed.
- More samples needed = $ higher costs $.
- Cost-effective management:
  - ***Don't*** assign 95% confidence levels by rote to all environmental analytical data. What do you really need?
  - Work with stakeholders and develop a logical plan.
  - Lower levels may be okay with initial "screening" data or with less harmful pollutants or with sites where use will be for light industry.

Determine Measurement Performance Criteria Carefully

- **Measurement performance criteria may consist of MQOs such as sensitivity, accuracy and precision.**
  - **Consider them for each pollutant and for each matrix.**
  - **Include decisions that will be made using the data.**
  - **Include the consequences of wrong decisions.**

- **Measurement performance criteria are often determined by the use of QA/QC samples.**
  - **QA/QC samples represent the quality of the data in the environmental samples**
  - **They measure the quality of the measuring system.**

Measurement Performance Criteria

Selecting the Correct QA/QC Samples is a Critical Component of Systematic Planning and Project Management for Field Studies

Don’t do it by mechanically !

Background Information for Putting QA/QC Into Perspective

What you will learn

- **Differences between QA and QC.**
- **How different types of QA/QC samples are used to measure different sources of error.**
  - **There are lots of ways to make mistakes**
  - **Do you want to know (1) if it happened, and (2) where it happened so the problem(s) can be fixed?**
- **What *bias* and *precision* are and how they affect the results of analytical data.**
- **How interferences can cause error in the data.**

Quality Assurance (QA)

Definition: **QA is a system of activities that assures the producer, user, or service that <u>defined standards of quality</u> with a <u>stated level of confidence</u> are met for a product.**

- **Source: John K. Taylor in *"Quality Assurance of Chemical Measurements,"* Lewis Publishers, Inc., 1987.**

Quality Control (QC)

Definition: **QC is a system of activities that *controls the quality* of a product or service so that it meets the needs of users.**

Q: Who specifies what QC is required with environmental data needed by a project?

A: Sometimes the methods used, sometimes the regulatory authority, sometimes the project manager.

What is the Difference Between QA and QC?

- **Thus QC consists of internal (technical) activities such as the use of QC samples, etc., to *control and assess the quality* of measurements while QA is the *management system* that ensures that an effective QC system is in place and working as intended.**
  - **Source: L. H. Keith in *Environmental Sampling and Analysis - A Practical Guide, Lewis Publishers, Inc., 1991***

Purpose of Quality Control

- **Purpose:** **The purpose of QC is to specify which activities must be conducted in order to assure that the data will meet QA criteria.**

- **Examples of QC activities: Specifying <u>what kinds?</u> and <u>how many?</u> QC samples needed.**
  - **Two separate questions involving different needs.**
  - **e.g. trip blanks, laboratory blanks, replicate samples, replicate analyses, laboratory control samples, field blanks, field spikes, laboratory spikes, glassware blanks, sampling equipment blanks, calibration curves, internal standards, external standards, standard addition, and many more.**
  - **<u>Each QC sample above is designed for a different and specific purpose to assess data or a process.</u>**

Measuring Sample Representativeness

- **Environmental samples are analyzed to form conclusions about a particular site.**
  - **Because you don't analyze the entire site you can only take samples that are <u>representative</u> of that site and analyze them.**
  - **You make conclusions on the extent of pollution at a site based on the evidence obtained from those representative samples.**
  - **What percent of the site do those samples represent? (often less than 0.001%)**
  - **What happens if those samples are not representative? (Wrong conclusions?)**

Why are QA/QC Samples Analyzed?

- QC samples are analyzed to form conclusions about the system that measures the kinds and amount of pollution in the environmental samples.

- IF there are problems with any of the QA/QC samples THEN there is doubt about the quality (reliability, accuracy, "goodness") of the environmental sample data. (Isn't this a decision rule form?)

- Sometimes the reason for analyzing QC samples is not well understood (i.e., they are done mechanically).
  - When this happens then if the QC data is useful it is only by good luck (and you can't depend on good luck!)

Measuring Heterogeneity

- **Target analytes are usually very unevenly distributed in environmental samples (including water and air) – this is "*heterogeneity*" or "*un-homogeneity*".**
  - **Measuring heterogeneity can provide information on how relevant the data from only a few samples is.**
  - **Analyze *co-located* samples (same area at a site separated by only a few meters from each other or, with fluid media, separated by only a few minutes or hours).**
    - **However, if many samples from a site will be analyzed then heterogeneity measurements may be of lesser importance.**

Measuring Analyte Migration

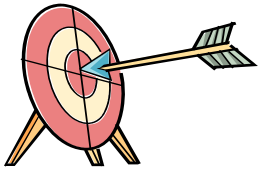
- **Pollutants migrate (*move*) –**
  - **Migration may be slow (PCBs down through a clay soil)**
  - **Relatively fast (an oil slick down a river)**
  - **Very fast (sulfur dioxide in ambient air from a stack plume)**

- **Using tracer compounds in QC samples introduced into the environmental medium to be sampled is one way to measure migration.**
  - **Another is to measure the movement of a unique chemical pollutant using its detection and/or concentration within a spatial area over time.**
  - **Each environmental project is unique.**

There Are Many Kinds of QA/QC Samples for Bias & Precision

- **There are samples that check for bias, i.e., *inaccuracy* (variance from the "true"answer).**
  - **Bias from procedures and bias from contamination.**
  - **Bias in sampling, bias in analysis, overall bias.**

- **There are samples that check for *imprecision* (variability from repeat samples or analyses).**

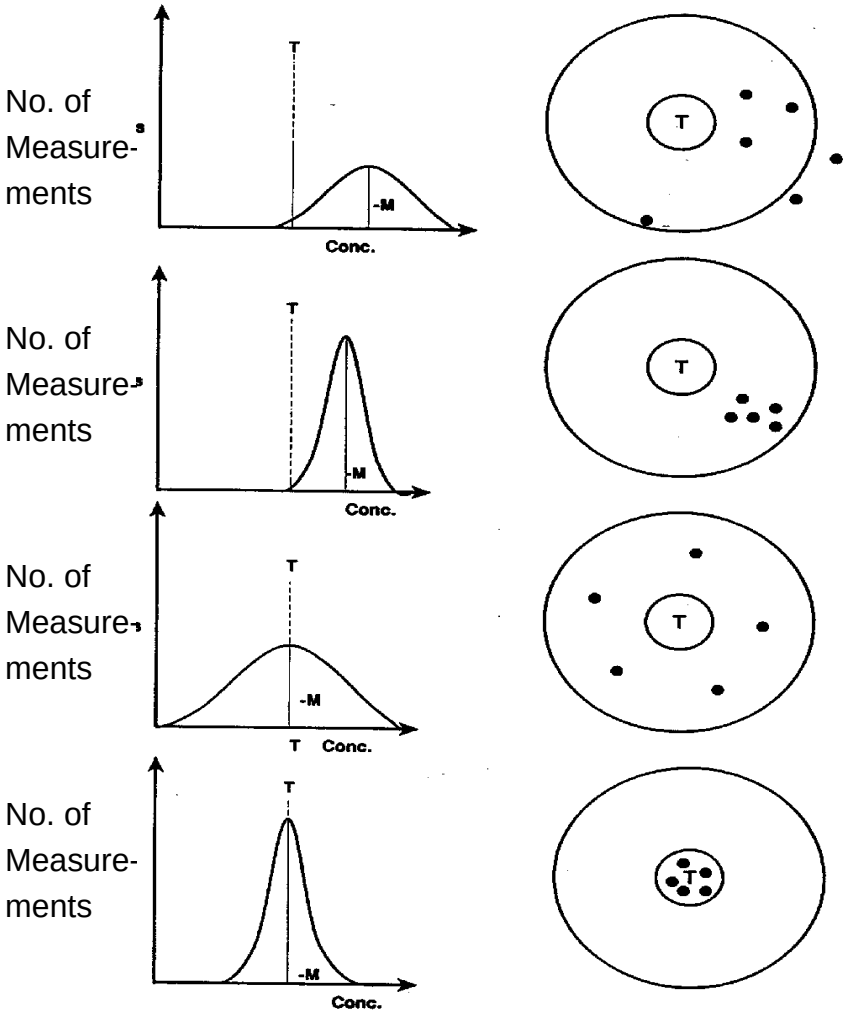


  - **Precision within one day or between many days.**
  - **Precision within one lab or between multiple labs.**
  - **Precision in sampling, precision in analysis, overall precision.**

Visual Representations of Precision and Accuracy

The four combinations of precision and bias (accuracy) are often represented as points on a bullseye target.

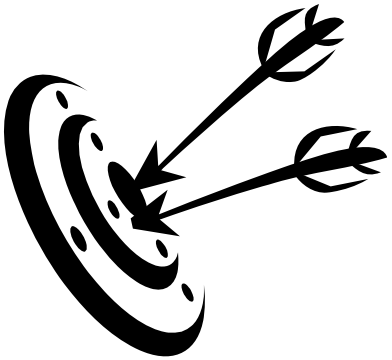
However, they can also be represented in terms of a Gaussian distribution of data around the true or assigned value.



CSP CHEMICAL SECURITY ENGAGEMENT PROGRAM
SAFETY AND SECURITY TRAINING

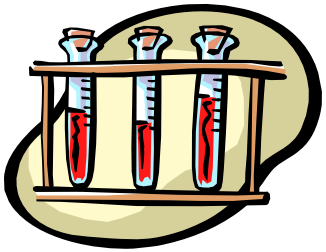
Characteristics of Precision

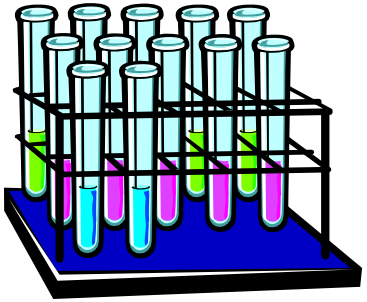
- **Precision is characterized by *random error*.**
  - **Sometimes the measured value will be higher than the true value and sometimes lower than the true value. It is random in nature.**
  - **If you measure many samples the errors caused by imprecision tend to "cancel out" and the average tends to be closer to the true value.**

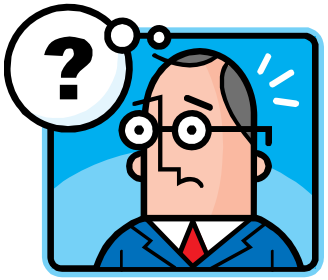


Characteristics of Bias

- **Bias (which results in inaccuracy – the "wrong" answer) is characterized by *systematic error*.**
  - **The error is always in the same direction; it causes inaccuracy to be more or less than the true answer.**
  - **Thus, no matter how many samples you analyze you will *never* get closer to the true answer.**



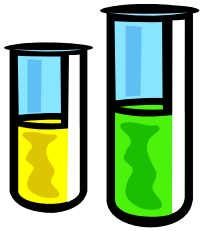




Interferences Are One Cause of Bias

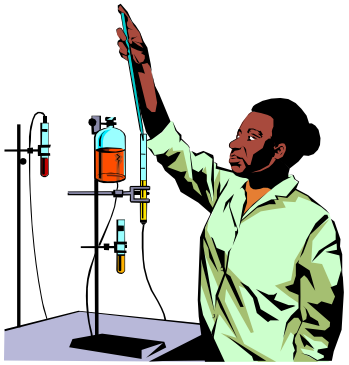
- A good example of a source of bias is an interference in the analytical method.

- There are two kinds of interferences:



  – 1 Those that are *inherent* in the samples, and

  – 2 Those caused by *procedures* in sampling & analysis



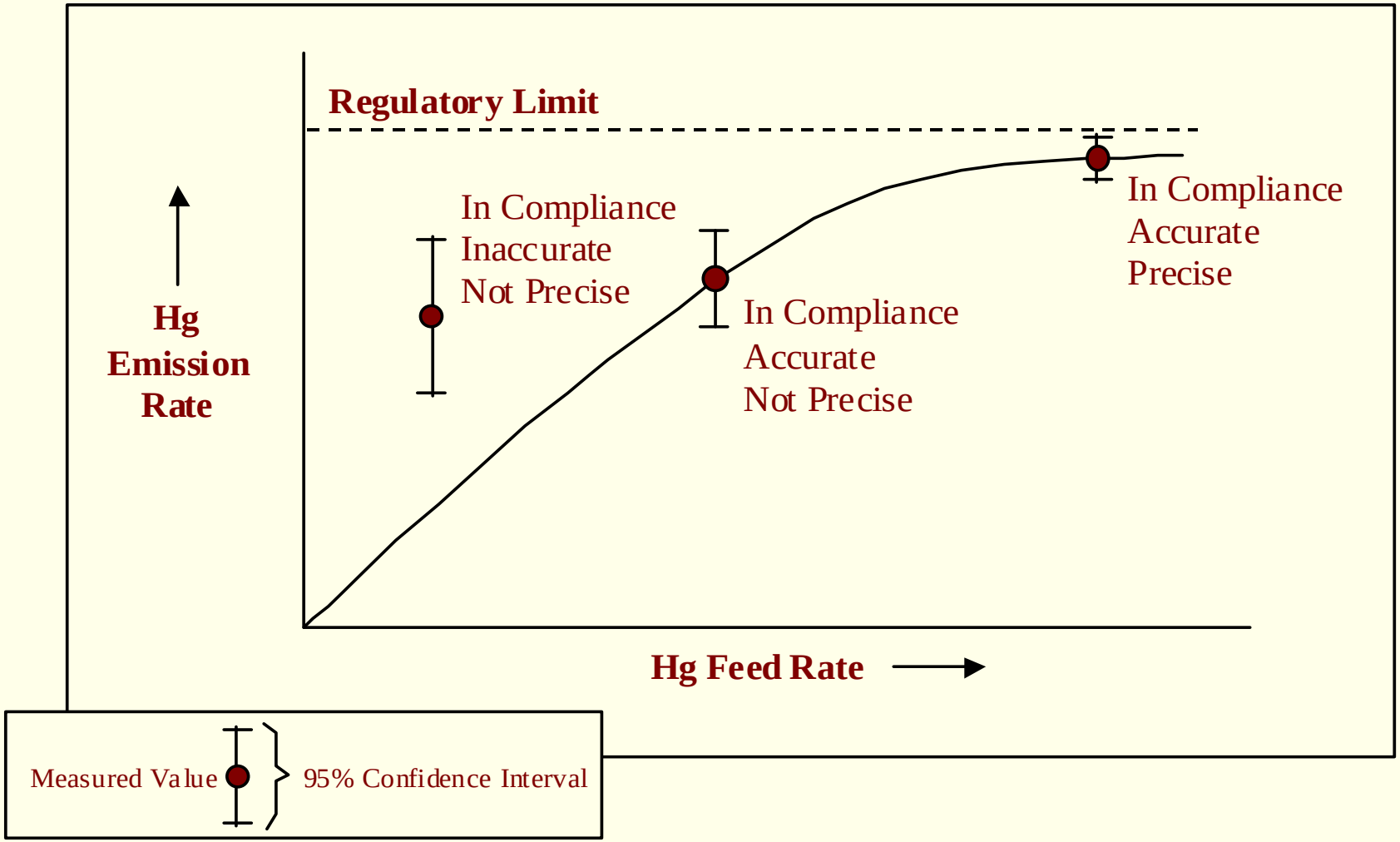
Inherent Interferences

- ***Inherent** interferences are part of the sample.*
  - **Characteristics of inherent interferences:**
    - **Can not be compensated for using the method.**
    - **Require modification or changing of the method employed OR require resampling and reanalysis with a different method after their source has been identified.**
    - **Their importance is that they may mask the signal of the pollutants (causing low values ↓ – *False Negatives*) OR they may add to it (causing higher than true values ↑ - *False Positives*).**
    - **Thus, their presence causes systematic bias - <u>always</u> producing either higher or lower than true values.**

Procedural Interferences

- ***Procedural** interferences are those that may be **added** during sampling, storage, sample preparation or analysis.*
  - **Characteristics of procedural interferences:**
    - **Can be removed and the existing method used again.**
    - **Importance is that they may mask the signal of the pollutants (causing low values ↓) or they may add to it (causing higher than true values ↑).**
    - **Thus, their presence causes systematic bias - *always* producing either higher or lower than true values.**

Accuracy and Precision Requirements as a Function of Regulatory Action Level



Summary

- **Systematic planning isn't as much fun as starting work on a project in the field.**

- **It may seem difficult to justify significant amounts of a budget on planning.**

- **BUT when projects exceed budgets and schedules because of unexpected results or because additional data is required by the stakeholders (e.g., regulators or community leaders), the source is almost always due to *insufficient planning*.**

- **Remember, keep plans detailed but flexible!**

DQO-PRO Tutorial- Introduction

Dr. Lawrence H. Keith President
Instant Reference Sources, Inc.

- This Tutorial is based on the DQO-PRO Version 2

- The program was written by Dr. Keith and two other scientists (G.L. Patton and P.G. Edwards) at Radian International LLC for the U.S. Government.

- DQO is free and may be copied and freely distributed.

- (Basis for Equations are found on the HELP tab of DQO-PRO Main Page)

DQO-PRO

- ***DQO-PRO*** is a series of statistical equations in the form of a simple calculator interface.
  - They can be used to help plan the minimum number of samples that should be collected in order to resolve three possible objectives in studies that require environmental sampling and analysis in order to make decisions.

The Purpose of DQO-PRO

- The statistical equations can also be used to assess the sufficiency of existing data to resolve decisions after sampling and analyses have taken place.
- Their purpose is to help optimize sampling designs (*make them the best*) so that objectives are met the first time and decisions are defensible to all relevant parties (for example, clients, regulatory agencies, etc).

The DQO-Process

- Remember, there are seven steps in the Data Quality Objective (DQO) Process. These are reviewed in the second tutorial module.
  - The **6th** step determines *how many samples* are required to meet your objectives based on the desired confidence levels (% probability) that you need.
  - Iteration (*repeating*) of this step is usually necessary in order to optimize the number of samples required to meet DQOs and meet budget, time and other study requirements.

Remember the Seven Stages of DQO Planning

1. State the Problem to be Resolved
2. Identify the Decision to be Made
3. Identify All the Inputs to the Decision
4. Narrow the Boundaries of the Study
5. Develop a Decision Rule
6. Develop Uncertainty Constraints
7. Optimize the Design for Obtaining Data

DQO-PRO Makes Calculations Easy

- The **6th** step determines **how many** samples are required to meet your objectives.
  - Iteration of this step is usually necessary in order to optimize the number of samples required to meet DQOs and meet budget, time and other study requirements.
  - ***DQO-PRO*** was developed to help make what are normally manual statistical calculation iterations *easy* to perform. It focuses on three common study objectives out of an almost infinite number of them**.**

DQO-PRO Background

- Programmed using Windows Visual Basic, ***DQO-PRO*** is accessed using Microsoft Windows.

- ***DQO-PRO Version 2.0*****,** which cost $15,000 to develop and update, can be freely copied and distributed**.**

DQO-PRO Use

- ***DQO-PRO*** provides answers for three common scenarios:
  1. Determining the **rate** at which an event occurs,
  2. Determining an estimate of an **average** within a tolerable error, and
  3. Determining the **sampling grid** necessary to detect "hot spots."
- The three associated tutorials show how easy it is to use ***DQO-PRO Version 2*** to estimate the *numbers of samples* you need for each of the above three scenarios based on <u>your personal</u> Data Quality Objectives.

One Thing to Remember!

Remember, information on the equations, assumptions used with the equations, and references to the equations and these assumptions are available from the available *DQO–PRO* program by clicking on

HELP

which is at the top of the program screen.

How to Contact the Author

Dr. Lawrence H. Keith, Monroe, GA USA

- **FAX: 1-253-595-7268**
- **e-mail: larrykeith@earthlink.net**
- **Internet: http://www.ChemistsHelpingChemists.org**

How to Determine How Many Samples to Analyze

Too Many and You Waste Money.

Too Few and You Waste Money Resampling,

or you Obtain Data of Insufficient Quality.

Basic Sampling Types

- **There are many different types of sampling**

  (stratified sampling, random sampling at sites judgmentally chosen etc).

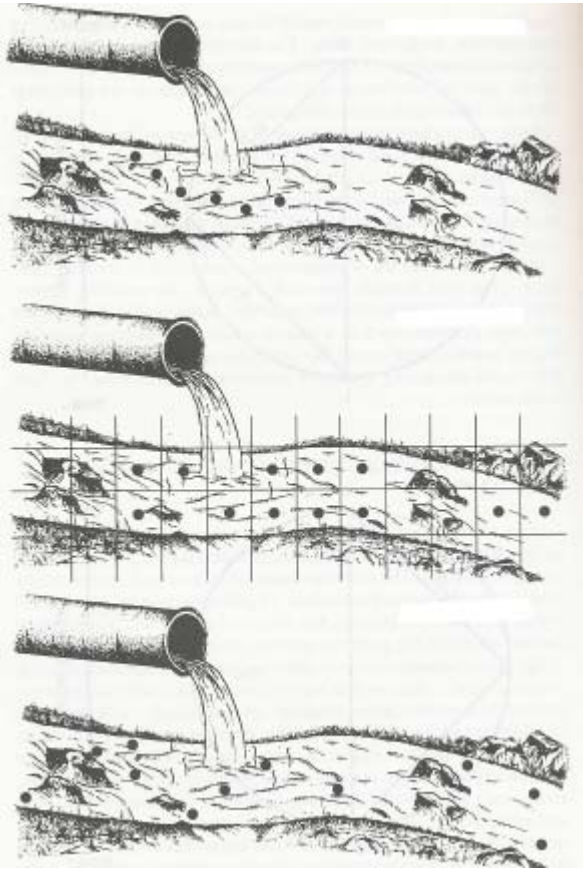
- **But they are all combinations or variations of three basic types of sampling:**

  1. Systematic Sampling

  2. Random Sampling,

  3. Judgmental Sampling

Representative sampling is a great challenge- sampling types



Judgmental Sampling

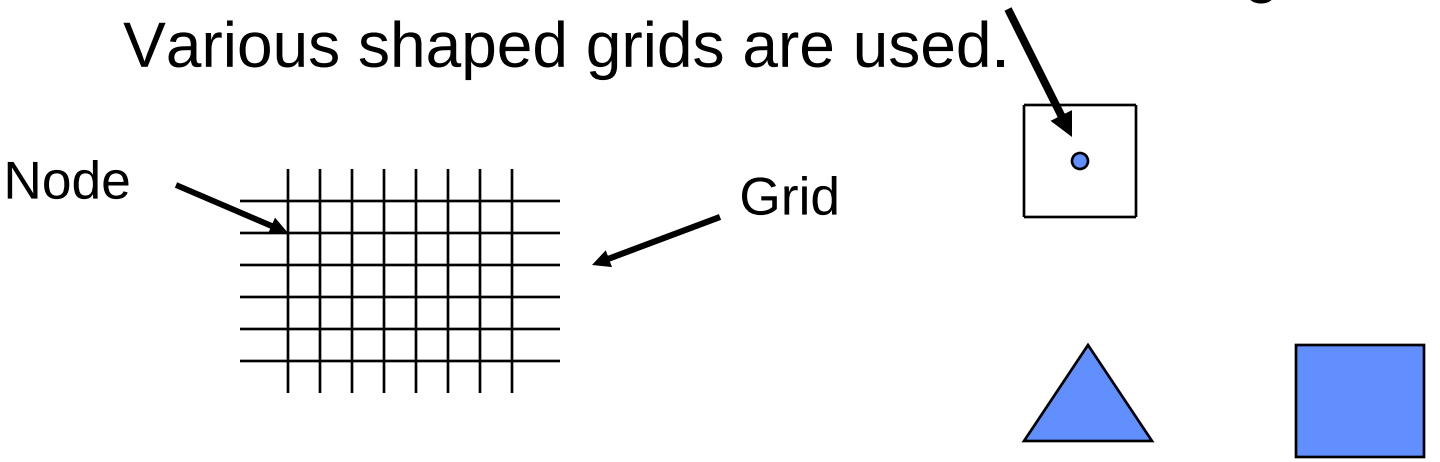
Systematic Sampling

Random Sampling

From: *Environmental Sampling and Analysis.* 1991 Larry Keith

Systematic Sampling

- ***Systematic sampling*** typically involves placing a ***grid*** over a map, or marking the sampling site, and then collecting samples at every ***node*** where the lines cross ***or*** in the ***center*** of each grid. Various shaped grids are used.



- The most common shapes are triangles, squares, or rectangles

Systematic Sampling

- Systematic sampling using a grid is usually used with contaminated land-based sites.

- It may be used to collect samples of
  - soil from a landfill,
  - to drill wells
  - to collect samples of groundwater, or
  - to collect aqueous sediments from the bottom of a lake or the sea.

Numbers of Samples and Systematic Planning

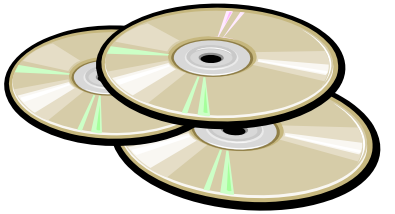
- **Numbers of samples depend on DQOs and MQOs when probabilistic sampling is performed.**

- **Remember, DQOs define the confidence *(% probability of being correct)* required in conclusions drawn from data.**
  - **You specify the amount of certainty required.**
  - **It should be part of sampling & analysis plan.**

- **MQOs are specific values that are used to measure the quality of analytical data.**
  - **Examples include recovery %, detection levels in ug/liter, etc.**

Numbers of Samples - Software to Calculate Them

- **High confidence levels (for example 95%) = more samples than lower confidence levels (for example 50%).**
  - **Numbers of samples are statistically calculated.**
  - **Confidence levels also are statistically calculated.**
  - **Equations are complex for people not skilled in statistical calculations.**



- **Free *DQO-PRO* software provides quick estimates of numbers of samples needed for a desired confidence level.**

DQO-PRO Software is One Tool to Calculate Data Quality

- **Programmed using Windows Visual Basic, *DQO-PRO* is accessed using Microsoft Windows**

- **It may be downloaded from the ACS Division of Environmental Chemistry home page which links to Dr. Keith's download page at:**

  http://www.instantref.com/dqo-form.htm

DQO-PRO Use

- *DQO-PRO* **provides answers for three common scenarios:**

1. **Determining the rate at which an event occurs,**

2. **Determining an estimate of an average within a tolerable error, and**

3. **Determining the sampling grid necessary to detect "hot spots".**

Components of DQO-PRO

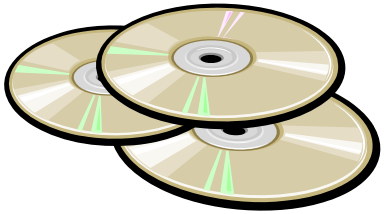
- ***HotSpot-Calc*** **- grid size and numbers of samples required at a land site to not miss a spot of contamination at a given confidence level.**

- ***Enviro-Calc*** **- number of samples needed for average concentration within a desired tolerable error at a given confidence level.**

- ***Success-Calc*** **- estimates rate at which an event occurs (e.g., the rates of false positives and false negatives at a given confidence level).**

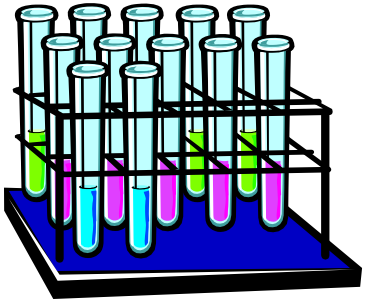
Enviro-Calc

DQO-PRO



Analytical Methods May Affect Sampling Plans

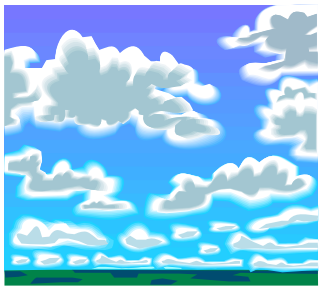
- **Ways in which sampling can affect the analysis of environmental samples are typically direct and usually very clear.**
- **However, it is the relationship in the opposite direction that can influence the number of environmental samples needed.**

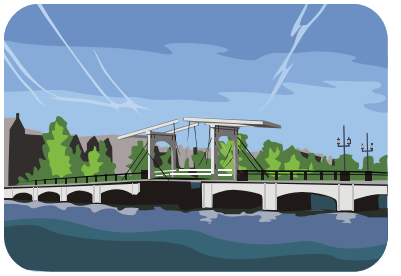


Calculating Numbers of Environmental Samples Needed

- **Statistical equations may be used when the purpose of environmental sampling is to estimate the average concentration of target analyte(s) within the study boundaries using random sampling.**

- **These types of samples are often used for fluid matrices such as air and water.**



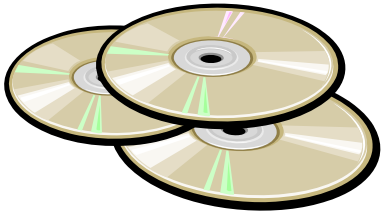


Advantages of *Enviro-Calc*

- **The great advantage lies in your ability to make *rapid assessments* of DQOs, MQOs, or project objective changes.**
  - **Often initial objectives cannot be met due to budget or time constraints, inability of the method to perform at its published levels (matrix effects, interferences, etc.).**
  - **You can quickly assess (either before or after samples have been analyzed) the consequences of missed DQOs and**
    - **Determine the consequences to either the maximum tolerable error or what the confidence level would be if fewer samples were analyzed.**

Success-Calc

DQO-PRO



QA and QC Samples Evaluate Environmental Data

- **QA and QC samples are used to provide an assessment of the kinds and amounts of bias and/or imprecision in the data obtained from environmental samples.**

- ***Thus, QA and QC samples are used to assess the collection and measurement system in a similar way that environmental samples are used to assess the portion of the environment from which they come*.**

- **<u>This is a key concept!</u>**

Environmental Data is Used to Evaluate Environmental Sites

- **Therefore, *representative* environmental samples are collected and analyzed to form conclusions about a particular site, and representative QC samples are analyzed to form conclusions about <u>the system that measures the environmental samples.</u>**

  - **This similarity in environmental sample usage and QC sample usage is often not appreciated or even recognized. *<u>Remember it.</u>***

What are Some Basic MQOs?

- **What percentage of false positives and false negatives are you willing to accept?**
  - **Most other factors considered will affect the amounts of false positives & false negatives.**
    - **Analyte Recovery**
    - **Detection Levels**
    - **Standards and Calibration Curves**
    - **Method Specificity**
    - **Interferences**

False Positives / False Negatives

"TRUTH"

SAMPLE CONCLUSION	PRESENT	ABSENT
PRESENT		False – Positive (Type I) Error
ABSENT	False – Negative (Type II) Error	

Method Blanks and Spikes

- **Method Blanks are used to demonstrate the absence of <u>False Positives</u> in analytical data.**
  - **These are samples (like a solvent) where the target analyte is known to *<u>NOT</u>* be present.**
- **Method Spikes (also called matrix spikes or laboratory control samples) are used to demonstrate the absence of <u>False Negatives</u> in analytical data.**
  - **These are samples where the target analyte is in *very small known amounts* in the sample matrix**

Beware of Prescriptive QA/QC

- **Assumption: The environmental samples analyzed in conjunction with a method blank and one or two spiked method blanks (or matrix spikes) will not result in false positive or false negative conclusions from the data because the accompanying *one or two* QA or QC samples did not contain them.**
  - **Not necessarily correct.**

- **The present way of assessing QA and QC data contains a basic flaw that is not usually recognized.**

DQO & MQO Failures Can Be Quantified

- **Numbers of QC samples needed to assure that a desired rate of false positive or negative conclusions would be detected could be compared with the number actually analyzed to quantify how well objectives were met.**

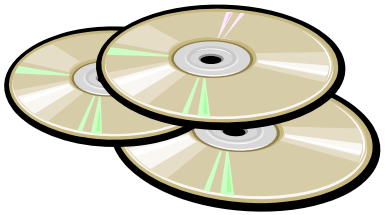
- **Furthermore, if the actual number of QC samples is less than the required number to meet the DQOs, then the increase in rate or decrease in probability of detecting false positives or false negatives (or any other specified characteristic) can be calculated.**

Large Numbers of QC Samples May Be Available From Historical Data

- ***Important*: many of the QC or environmental samples needed for a statistical population are available (or can easily be made available) – they are just not used this way now.**
- **Thus, increased costs associated with large numbers of samples may not be necessary – they may be minimal or even reduced with proper planning.**
- **How is proper planning done? For example, by obtaining historical QC sample data from the lab under the same conditions that the test samples were obtained within a reasonable window of time that covers the time the samples were analyzed.**

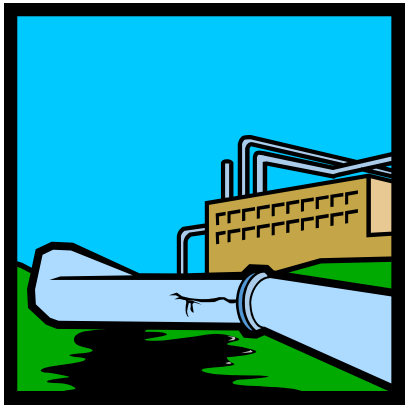
HotSpot-Calc

DQO-PRO



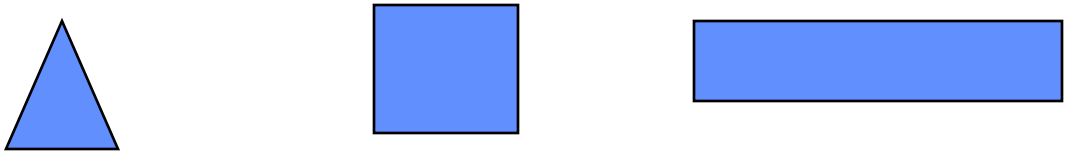
Hot Spot Sampling Objectives

- **The primary objective of "hot spot" sampling is to determine if localized areas of contamination exist.**
  - **May be due to spills, leaks, buried waste, or any number of other events where contamination might be confined to a relatively small area.**
  - **A single site might have multiple hot spots of different origins.**



Systematic Sampling

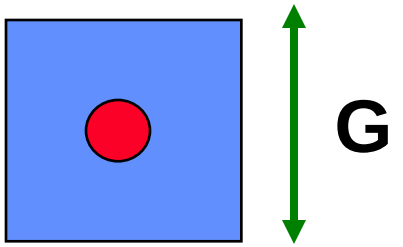
- Remember, it typically involves placing a grid over a map, or marking the sampling site, and then collecting samples at <u>every node</u> where the lines cross or in the center of each grid.
  - Various shaped grids are used.
  - The most common shapes for grids are triangles, squares, or rectangles.



Calculating Numbers of Samples

The number of samples required for hot spot sampling is the number of samples required to sample all grid areas at the site for the selected grid spacing. The number of samples required for a square grid is approximated by the equation:

$$n = A/G^2$$



where,

n = number of samples,

A = area to be sampled, in the square of the units for G

G = grid spacing

HotSpot-Calc

- ***HotSpot-Calc*** **is designed to determine the grid spacing needed to detect the presence of a single hot spot.**
  - **Spot of specified size and shape**
  - **Includes specified probability of <u>missing</u> the hot spot.**
  - **It is based on the following key assumptions:**

1. **That the hot spot is circular, short elliptical or long elliptical in shape; and**
2. **That sample measurements are collected on square, rectangular, or triangular grids.**

HotSpot-Calc Probability

- **The probability (% chance) of finding a hot spot is…**
  - **A function of the specified size and shape of the hot spot, the pattern of the grid, and the relationship between the size of the hot spot and the grid spacing.**

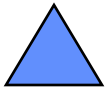
- ***HotSpot-Calc* is programmed based on the probability of missing a hot spot if one exists rather than on the probability of finding one. (Not a problem!)**
  - **Answer is based on the statistical equations used and the perspective of the Scientific Method question**
    - **Assumes that a hot spot is present (easiest assumption).** ***Therefore, if a hot spot is present, what is the probability of missing it?***

The Effects of Grid Shapes

1. The number of samples needed <u>increases</u> as the size of the spot which is acceptable to miss <u>decreases.</u>

2. The number of samples needed <u>increases</u> as the acceptable probability of missing a hot spot <u>decreases</u>,

3. If the hot spot is circular fewer numbers of samples are needed than when it is elliptical and, the longer the horizontal axis is in the ellipse the larger is the number of samples that will be needed for a given probability and grid shape.



4. A triangular grid is the *most* efficient and a rectangular grid is the *least* efficient for finding a hot spot using the same assumptions.

The Benefits of Using Hot-Spot Calc

1. If you are concerned with the **probability of missing** a spot of a certain minimum size then you'll know how many samples to take over the sampling area to not miss a spot of that size if it exists.

2. If you are constrained by budget and can only collect a certain number of samples at a site then you'll know how **large the spot** of contamination would have to be for you to ***not*** miss it ***IF*** it is present.

Here is a Useful Performance Criteria Check List

- [ ] Have all initial analytes been documented and agreed to by all the stakeholders?
- [ ] Have all the initial media been selected for analysis and agreed to by the stakeholders?
- [ ] Have all the action levels for each analyte in each media or matrix been agreed to?
- [ ] Have all the confidence levels for each analyte in each media or matrix been agreed to?
- [ ] Do all the selected methods appear capable of meeting QC performance criteria?
- [ ] Do planned analyses meet time and cost budgets?