Principles of environmental sampling

An integral part of environmental sampling is the design of the sampling program.. This section will emphasise the integration between good design principles and aspects of sampling involving choice of parameters and methods.

A useful reference for all you want to know about sampling is **Keith (1996)**. He proposes a structured approach to environmental sampling:

Determining Number and Kinds of Samples				
Types of samples	**Environmental samples**		**QC samples**	
	What kinds?	**How many?**	**What kinds?**	**How many?**
Principles	Environmental samples must be representative of the portion of the environment being investigated	Procedures for sampling and analysis influence each other, so plans for sampling and analysis are co-dependent	QC samples must be representative of the environmental samples being analysed	QC samples are used to provide an assessment of the kinds and amounts of bias and imprecision in data from analysis of the environmental samples.

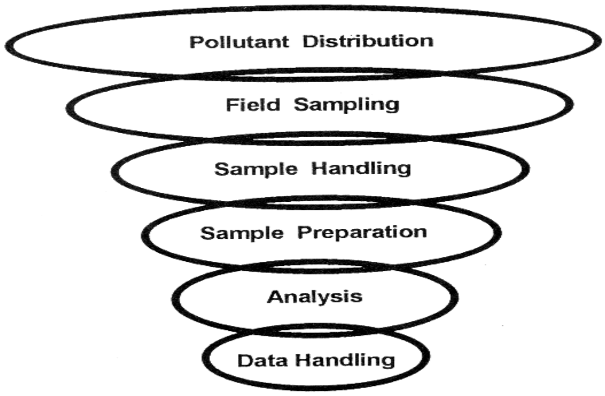
Environmental samples must be representative of the portion of the environment being investigated.

The point of sampling, and statistical analysis, is to use a small number of samples to represent the characteristics of the whole population that make up the area (and/or times) of interest. The characteristics of the site must be understood and appropriately sampled if that subset is to be representative.

Both the parameters of concern and the impacts being investigated can partition differently into different matrices. The primary matrices include: water; soils; biota; air; solid wastes; and liquid wastes. Each of these primary matrices can be subdivided and many parameters will dynamically partition between matrices.

Practical considerations of sampling, materials handling, storage, and laboratory procedures need to be considered in determining the most appropriate protocols to ensure that the samples will be representative.

“The activities associated with collecting and preserving environmental samples are generally acknowledged to be the weakest link in the chain of activities that comprise environmental sampling and analysis. Sampling activities are often the weakest link because they usually contribute the largest amount of errors in conclusions about the ‘true’ condition of the site being investigated. [The figure below] illustrates the relative sources of variability in the chain of typical sampling and analysis events.”



Sources of variability in the sampling and analysis chain of events.

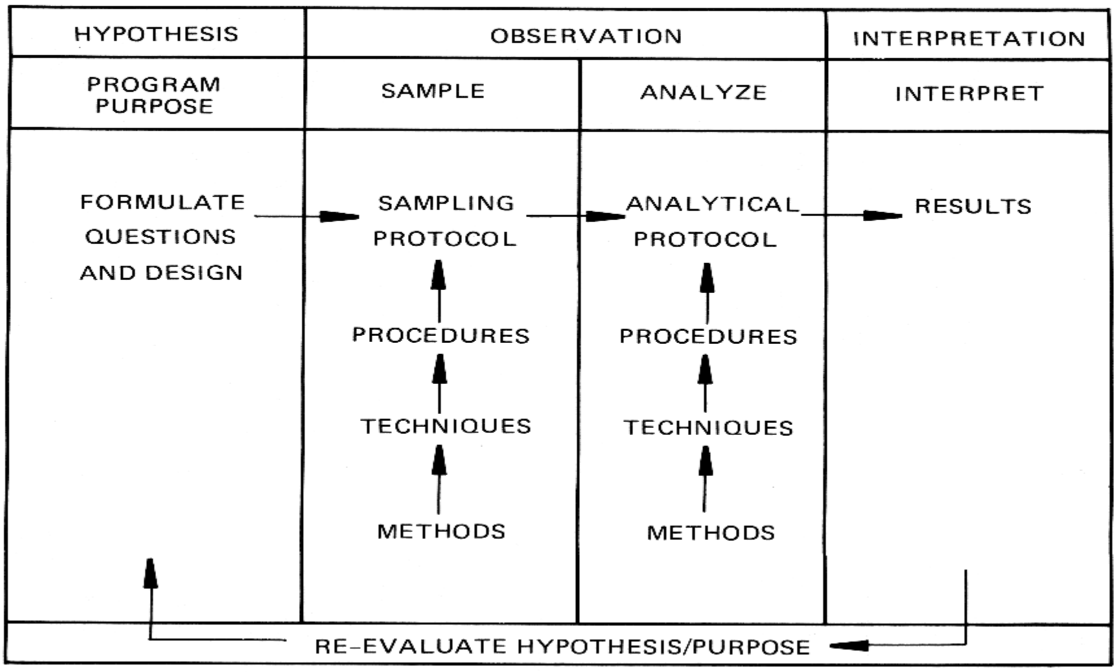
Although laboratory and data handling variability are often best controlled they can receive the most attention in QC efforts. The design of the program and the consequent field sampling are the areas most difficult to control and supervise. Despite the lack of impressive equipment field sampling can require the greatest levels of skill and experience and, conversely, good field workers should be included in the analytical loop so they get feedback on their efforts.

Procedures for sampling and analysis influence each other, and so plans for sampling and analysis are co-dependent.

The issues associated with power, or the avoidance of Type II errors is discussed elseshere. The variables that influence power are directly associated with sampling and need to be considered at the same time:

- α, the acceptable error rate for not detecting the parameter (a Type I error)
- the **variance** in the test to detect the actual population
- the ability to detect the predetermined **effect level** that would require action
- the **number of samples** to be collected
- confidence that the effect has been detected (equivalent to $1-\beta$, the **power** of the test).

Given any four the other can be **calculated.** Clearly, practical considerations of sample size, handling, methods, costs, etc are directly linked to the ability to control variance, α, β, and *effect size*. Using a less expensive sampling protocol of low precision (without collecting more samples) which yields unusable results is not good economy. Conversely, linking an expensive and highly precise analysis to exploratory field surveys could lead to misplaced confidence and wasted resources.



The key issue is that the data obtained can be used with the specified level of confidence to make an inference about the environmental population from which samples were collected and analysed. This includes:

- The level of confidence in the correct identification of the analyte(s)/indicator(s) etc
- The concentration levels at which the analyte(s) can be identified or quantified with some specified level of confidence
- The degree to which the environmental samples represent the remaining unsampled population at an environmental site.

The US EPA has developed the **Data Quality Objectives** (DQOs) process to improve environmental sampling. The process requires the investigator:

1. State the problem to be resolved
2. Identify the decision(s) that must be made
3. Identify all the inputs needed to make the decision(s)
4. Define the study boundaries (e.g. in space, time, and analytes)
5. Develop a decision rule for each decision
6. Specify limit on decision errors, and
7. Optimise the design for collecting data

These steps should all be included in a **Error! Hyperlink reference not valid.**, taking into account both sampling and analysis.

QC samples must be representative of the environmental samples being analysed.

QC samples have the same relationship to the sample population as the sample population to the total environmental sample. Thus, QC samples are used to assess the collection and measurement system in a manner similar to the way that environmental samples are used to assess the portion of the environment site from which they come. The intention, in both cases, is that valid inferences can be drawn about the larger population, particularly in regard to **the bias and precision** in the sample populations.

QC samples are used to provide an assessment of the kinds and amounts of bias and imprecision in data from analysis of the environmental samples.

In the same way that experimental design needs to address the issues of false negatives (Type II errors) and false positives (Type I errors) QC samples assess the same issues to ensure that the analytical results are reliable. False positives can arise from contamination or misidentification of other analytes; false negatives commonly arise from poor recovery (and/or preservation) of analytes or interference from other compounds. A QC program needs to address all levels of concern in sample handling and testing.

As I discussed above for the environmental samples the QC samples also need to be defined in terms of the confidence limits as to false negatives and false positives. Using a single sample in a batch process is unlikely to give this information; achieving predetermined confidence limits will require information on the variance of the test and QC samples, the effect size to be discriminated, and the number of samples. This broader view is required so that the focus remains on the quality of conclusions based on the data rather than an overly technical focus on “data quality”.