Designation: D6044 – 21

Standard Guide for Representative Sampling for Management of Waste and Contaminated Media[1]

This standard is issued under the fixed designation D6044; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ε) indicates an editorial change since the last revision or reapproval.

1. Scope

1.1 This guide covers the definition of representativeness in environmental sampling, identifies sources that can affect representativeness (especially bias), and describes the attributes that a representative sample or a representative set of samples should possess. For convenience, the term "representative sample" is used in this guide to denote both a representative sample and a representative set of samples, unless otherwise qualified in the text.

1.2 This guide outlines a process by which a representative sample may be obtained from a population. The purpose of the representative sample is to provide information about a statistical parameter(s) (such as mean) of the population regarding some characteristic(s) (such as concentration) of its constituent(s) (such as lead). This process includes the following stages: (*1*) minimization of sampling bias and optimization of precision while taking the physical samples, (*2*) minimization of measurement bias and optimization of precision when analyzing the physical samples to obtain data, and (*3*) minimization of statistical bias when making inferences from the sample data to the population. While both bias and precision are covered in this guide, major emphasis is given to bias reduction.

1.3 This guide describes the attributes of a representative sample and presents a general methodology for obtaining representative samples. It does not, however, provide specific or comprehensive sampling procedures. It is the user's responsibility to ensure that proper and adequate procedures are used.

1.4 The assessment of the representativeness of a sample is not covered in this guide since it is not possible to ever know the true value of the population.

1.5 Since the purpose of each sampling event is unique, this guide does not attempt to give a step-by-step account of how to develop a sampling design that results in the collection of representative samples.

1.6 Appendix X1 contains two case studies which discuss the factors for obtaining representative samples.

1.7 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety, health, and environmental practices and determine the applicability of regulatory limitations prior to use.*

1.8 *This international standard was developed in accordance with internationally recognized principles on standardization established in the Decision on Principles for the Development of International Standards, Guides and Recommendations issued by the World Trade Organization Technical Barriers to Trade (TBT) Committee.*

2. Referenced Documents

2.1 *ASTM Standards:*[2]

D3370 Practices for Sampling Water from Flowing Process Streams
D4448 Guide for Sampling Ground-Water Monitoring Wells
D4547 Guide for Sampling Waste and Soils for Volatile Organic Compounds
D4823 Guide for Core Sampling Submerged, Unconsolidated Sediments
D5088 Practice for Decontamination of Field Equipment Used at Waste Sites
D5681 Terminology for Waste and Waste Management
D5792 Practice for Generation of Environmental Data Related to Waste Management Activities: Development of Data Quality Objectives
D5956 Guide for Sampling Strategies for Heterogeneous Wastes
D6051 Guide for Composite Sampling and Field Subsampling for Environmental Waste Management Activities
D6169 Guide for Selection of Soil and Rock Sampling Devices Used With Drill Rigs for Environmental Investigations
D6286 Guide for Selection of Drilling and Direct Push

---

[1] This guide is under the jurisdiction of ASTM Committee D34 on Waste Management and is the direct responsibility of Subcommittee D34.01.01 on Planning for Sampling.

Current edition approved Oct. 1, 2021. Published October 2021. Originally approved in 1996. Last previous edition approved in 2015 as D6044 – 96 (2015). DOI: 10.1520/D6044-21.

[2] For referenced ASTM standards, visit the ASTM website, www.astm.org, or contact ASTM Customer Service at service@astm.org. For *Annual Book of ASTM Standards* volume information, refer to the standard's Document Summary page on the ASTM website.

Methods for Geotechnical and Environmental Subsurface Site Characterization

D6634 Guide for Selection of Purging and Sampling Devices for Groundwater Monitoring Wells

D6771 Practice for Low-Flow Purging and Sampling for Wells and Devices Used for Ground-Water Quality Investigations

D7929 Guide for Selection of Passive Techniques for Sampling Groundwater Monitoring Wells

3. Terminology

3.1 *Definitions*—For definitions of terms used in this standard, refer to Terminology D5681.

3.2 *Definitions of Terms Specific to This Standard:*

3.2.1 *analytical unit, n*—the actual amount of the sample material analyzed in the laboratory.

3.2.2 *bias, n*—a systematic positive or negative deviation of the sample or estimated value from the true population value.

3.2.2.1 *Discussion*—This guide discusses three sources of bias: sampling bias, measurement bias, and statistical bias.

3.2.2.2 *Discussion*—There is a sampling bias when the value inherent in the physical samples is systematically different from what is inherent in the population.

3.2.2.3 *Discussion*—There is a measurement bias when the measurement process produces a sample value systematically different from that inherent in the sample itself, although the physical sample is itself unbiased. Measurement bias can also include any systematic difference between the original sample and the sample analyzed, when the analyzed sample may have been altered due to improper procedures such as improper sample preservation or preparation, or both.

3.2.2.4 *Discussion*—There is a statistical bias when, in the absence of sampling bias and measurement bias, the statistical procedure produces a biased estimate of the population value.

3.2.2.5 *Discussion*—Sampling bias is considered the most important factor affecting inference from the samples to the population.

3.2.3 *biased sampling, n*—the taking of a sample(s) with prior knowledge that the sampling result will be biased relative to the true value of the population.

3.2.3.1 *Discussion*—This is the taking of a sample(s) based on available information or knowledge, especially in terms of visible signs or knowledge of contamination. This kind of sampling is used to detect the presence of localized contamination or to identify the source of a contamination. The sampling results are not intended for generalization to the entire population. This is one form of authoritative sampling (see *judgment sampling*).

3.2.4 *composite sample, n*—a combination of two or more samples.

3.2.5 *constituent, n*—an element, component, or ingredient of the population.

3.2.5.1 *Discussion*—If a population contains several contaminants (such as acetone, lead, and chromium), these contaminants are called the constituents of the population.

3.2.6 *error, n*—the random or systematic deviation of the observed sample value from its true value (see *bias* and *sampling error*).

3.2.7 *homogeneity, n*—the condition of the population under which all items of the population are identical with respect to the characteristic(s) of interest.

3.2.8 *judgment sampling, n*—taking of a sample(s) based on judgment that it will more or less represent the average condition of the population.

3.2.8.1 *Discussion*—The sampling location(s) is selected because it is judged to be representative of the average condition of the population. It can be effective when the population is relatively homogeneous or when the professional judgment is good. It may or may not introduce bias. It is a useful sampling approach when precision is not a concern. This is one form of authoritative sampling (see *biased sampling*).

3.2.9 *representative sample, n*—a sample collected in such a manner that it reflects one or more characteristics of interest (as defined by the project objectives) of a population from which it is collected.

3.2.9.1 *Discussion*—A representative sample can be a single sample, a collection of samples, or one or more composite samples. A single sample can be representative only when the population is highly homogeneous.

3.2.10 *sample, n*—one or more items or portions collected from a lot or population.

3.2.10.1 *Discussion*—Sample is a term with numerous meanings. The scientist collecting physical samples (for example, from a landfill, drum, or monitoring well) or analyzing samples considers a sample to be that unit of the population that was collected and placed in a container. A statistician considers a sample to be a subset of the population, and this subset may consist of one or more physical samples. To minimize confusion, the term *sample*, as used in this guide, is a reference to either a physical sample held in a sample container, or that portion of the population that is subjected to in-situ measurements, or a set of physical samples. See *representative sample*.

3.2.10.2 *Discussion*—The term *sample size* also means different things to the scientist and the statistician. To avoid confusion, terms such as sample mass/sample volume and number of samples are used instead of sample size.

3.2.11 *sampling error, n*—the systematic and random deviations of the sample value from that of the population. The systematic error is the *sampling bias*. The random error is the *sampling variance*.

3.2.11.1 *Discussion*—Before the physical samples are taken, potential sampling variance comes from the inherent population heterogeneity (sometimes called the "fundamental error"; see *heterogeneity*). In the physical sampling stage, additional contributors to sampling variance include random errors in collecting the samples. After the samples are collected, another contributor is the random error in the measurement process. In each of these stages, systematic errors can occur as well, but they are the sources of bias, not sampling variance.

3.2.11.2 *Discussion*—Sampling variance is often used to refer to the total variance from the various sources.

3.2.12 *stratum, n*—a subgroup of the population separated in space or time, or both, from the remainder of the population, being internally similar with respect to a target characteristic of interest, and different from adjacent strata of the population.

3.2.12.1 *Discussion*—A landfill may display spatially separated strata, such as old cells containing different wastes than new cells. A waste pipe may discharge into temporally separated strata of different constituents or concentrations, or both, if night-shift production varies from the day shift. In this guide, strata refer mostly to the stratification in the concentrations of the same constituent(s).

4. Significance and Use

4.1 This guide defines the meaning of a representative sample, as well as the attributes the sample(s) needs to have in order to provide a valid inference from the sample data to the population.

4.2 This guide also provides a process to identify the sources of error (both systematic and random) so that an effort can be made to control or minimize these errors. These sources include sampling error, measurement error, and statistical bias.

4.3 When the objective is limited to the taking of a representative (physical) sample or a representative set of (physical) samples, only potential sampling errors need to be considered. When the objective is to make an inference from the sample data to the population, additional measurement error and statistical bias need to be considered.

4.4 This guide does not apply to the cases where the taking of a nonrepresentative sample(s) is prescribed by the study objective. In that case, sampling approaches such as judgment sampling or biased sampling can be taken. These approaches are not within the scope of this guide.

4.5 Following this guide does not guarantee that representative samples will be obtained. But failure to follow this guide will likely result in obtaining sample data that are either biased or imprecise, or both. Following this guide should increase the level of confidence in making the inference from the sample data to the population.

4.6 This guide can be used in conjunction with the DQO process (see Practice D5792).

4.7 This guide is intended for those who manage, design, and implement sampling and analytical plans for waste management and contaminated media.

5. Representative Samples

5.1 Samples are taken to make inferences about some statistical parameter(s) of the population regarding some characteristic(s) of its constituent(s) of interest. This is discussed in the following sections.

5.2 *Samples*—When a representative sample consists of a single physical sample, it is a sample that by itself reflects the characteristics of interest of the population. On the other hand, when a representative sample consists of a set of physical samples, the samples collectively reflect some characteristics of the population, though the samples individually may not be representative. In most cases, more than one physical sample is necessary to characterize the population, because the population in environmental sampling is usually heterogeneous.

5.3 *Constituents and Characteristics*—A population can possess many constituents, each with many characteristics. Usually it is only a subset of these constituents and characteristics that are of interest in the context of the stated problem. Therefore, samples need to be representative of the population only in terms of these constituent(s) and characteristic(s) of interest. A sampling plan needs to be designed accordingly.

5.4 *Parameters*—Similarly, samples need to be representative of the population only in the parameter(s) of interest. If the interest is only in estimating a parameter such as the population mean, then composite samples, when taken correctly, will not be biased and therefore constitute a representative sample (regarding bias) for that parameter. On the other hand, if the interest happens to be the estimation of the population variance (of individual sampling units), another parameter, then the variance of the composite samples is a biased estimate of the population variance and therefore is not representative. (It is to be noted that composite samples are often used to increase the precision in estimating the population mean and not to estimate the population variance of individual sampling units.)

5.5 *Population*—Since the samples are intended to be representative of a population, a population must be well defined, especially in its spatial or temporal boundaries, or both, according to the study objective.

5.6 *Representativeness*—The word "reflects" in this guide is used to mean a certain degree of low bias and high precision when comparing the sample value(s) to the population value(s). This is a broad definition of sample representativeness used in this guide. A narrower definition of representativeness is often used to mean simply the absence of bias.

5.6.1 *Bias*—Bias is sometimes mistakenly taken to be "a difference between the observed value of a physical sample and the true population value." The correct definition of bias is "a *systematic* (or consistent) difference between an observed (sample) value and the true population value." The word "systematic" here implies "on the average" over a set of physical samples, and not a single physical sample. Recall that sampling error consists of the random and systematic deviations of a sample (or estimated) value from that of the population. Although random deviations may occur on occasions due to imprecision in the sampling or measurement processes, or both, they balance out on the average and lead to no systematic difference between the sample (or estimated) value and the population value. The random deviation corresponds to the observation of "a random difference between a single physical sample value and the true population value," which can be randomly positive or negative, and is not a bias. On the other hand, a persistent positive or negative difference is a systematic error and is a bias.

5.6.1.1 In order to assess bias, the true population value must be known. Since the true population value is rarely known, bias cannot be quantitatively assessed. However, this guide provides an approach to identifying the potential sources of bias and general considerations for controlling or minimizing these potential biases.

5.6.2 *Precision*—Precision has to do with the level of confidence in estimating the population value using the sample data. If the population is totally homogeneous and the measurement process is flawless, a single sample will provide a completely precise estimate of the population value. When the population is heterogeneous or the measurement process is not totally precise, or both, a larger number of samples will provide a more precise estimate than a smaller number of samples.

5.6.2.1 In the case of bias, the goal in environmental sampling is its absence. In the case of precision, the goal in sampling will depend on factors such as:

(1) The precision level needed to achieve the desired levels of decision errors, both false positive and false negative errors,

(2) If the true value is known or suspected to be well below the regulatory limit, high precision in the samples may not be needed, and

(3) The study budget.

5.6.2.2 Note that the second item applies similarly to bias as well.

5.6.2.3 Since bias, especially during sampling, can be very large when proper procedures are not followed, it is considered to be the first necessary condition for sample representativeness. On the other hand, precision can be more or less controlled, for example, by increasing the number of samples taken or by decreasing the sampling or measurement variabilities, or both.

5.6.2.4 The optimal number of samples to take to achieve a desired level of precision is typically an issue in optimization of a sampling plan. Therefore, the precision issue will be covered only briefly in this guide.

6. A Systematic Approach to Representative Sampling

6.1 A systematic approach is one that first defines the desired end result and then designs a process by which such a result can be obtained. In representative sampling, the desired end result is a sample or a set of samples that achieves desired levels of low bias and high precision.

6.2 A representative sampling process is described in Fig. 1. The key components in the process are described in this section.

6.3 *Study Objective*—A sampling plan is designed according to a defined problem or a stated study objective. The samples are then collected according to the sampling plan. Generally, the study objective dictates that representative samples be taken for the purpose of inference about the population. In that case, these samples will need to be collected according to this guide in order for the inference to be valid. Occasionally, the objective is merely to detect the presence of a contaminant or to obtain a "worst case" sample. In that case, an authoritative sampling approach (biased sampling or judgment sampling) may be taken and this guide does not apply.

6.4 *Population*—A population consists of the totality of items or units of materials under consideration (Compilation of ASTM Standard Definitions, 1990). Its boundaries (spatial or temporal, or both) are defined according to the problem statement. This population is usually called the *target population*. In order to solve the stated problem, samples must be taken from the target population.

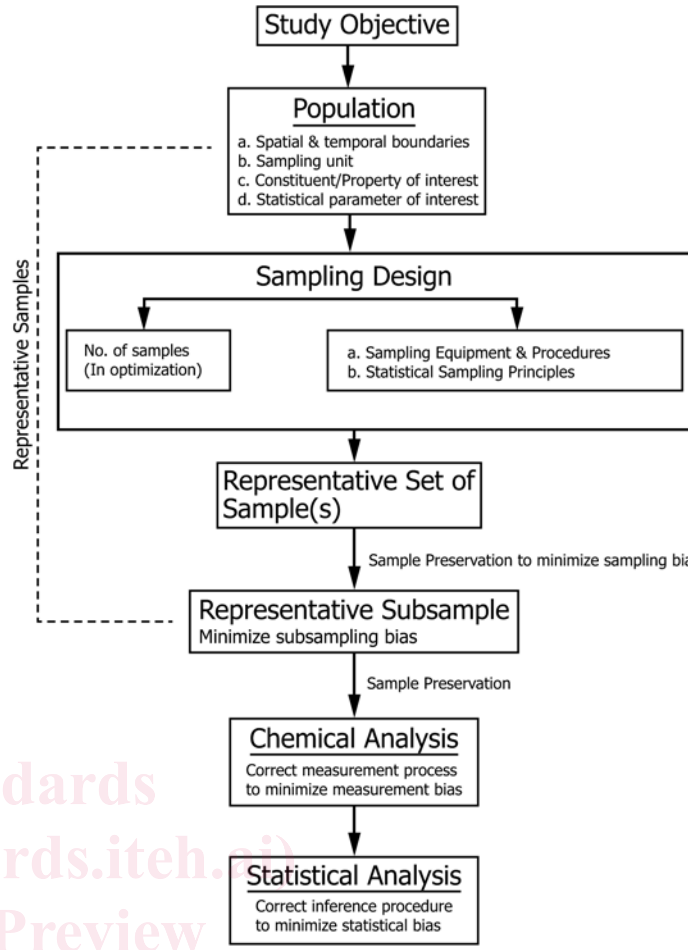

FIG. 1 A Systematic Approach to Representative Sampling

6.4.1 *Sampled Population*—Sometimes some parts of the target population may not be amenable to sampling due to factors such as accessibility. The boundaries of the target population actually sampled due to factors such as incomplete accessibility define the sampled population.

6.4.1.1 Although the samples taken from the sampled population may be representative of the sampled population, they may not be representative of the target population. In this case, potential exists that the samples taken from the sampled population may systematically deviate from the true value of the target population, thereby introducing bias when making inference from the samples to the target population.

6.4.1.2 When the boundaries of the target and sampled populations are not identical, some possible solutions are:

(1) The parties to the decision-making may agree that the sampled population is a sufficient approximation to the target population. A sampling plan can then be designed to take representative samples from the "sampled population";

(2) Qualifications on the sampling results are made based on the differences between the two populations. Some professional judgment may have to be exercised here; and

(3) Redefine the problem by considering what problem is solvable based on the observed differences between the two populations.

6.4.1.3 Occasionally, the sampled population is chosen on purpose to be different from the target population. For example, an investigator may be interested in the lead content in the sludge of a surface impoundment (the target population). He may decide to take samples from the sludge near the inlet (sampled population). Thus, the impoundment is the target population, while the inlet area is the sampled population. If the interest is in the target population, then this is an example of a biased sampling approach. On the other hand, the involved parties may decide to redefine the target population to include only the inlet area. Then the target population and the sampled population are identical. Again, the definition of a population depends on the problem statement.

6.4.1.4 In yet other circumstances, an investigator may take only a sample from the population. The following cases are possible:

(1) This one physical sample can be a sample from a biased sampling approach, for the purpose of detecting the presence of a contaminant or identifying the source of contamination. Therefore, it is not a representative sample due to its bias;

(2) This one physical sample can be a sample from judgment sampling, for the purpose of estimating the average condition of the population. Bias may or may not exist depending to some degree on the expertise of the sampler;

(3) This sample can be viewed as a population itself if the investigator is interested in the sample alone and a result from this sample is not to be used to infer to areas outside the sample. In this case, no bias exists; and

(4) If this sample is the composite of a few samples taken from the population, bias is likely to be minimal if the original samples are carefully taken.

6.4.2 *Decision Unit*—Often a population may be divided into several exposure units, cleanup units, or strata. If the environmental management decision is to be made for the entire population as a whole, representative samples can be obtained by designs such as a stratified random sampling design. Here the entire population is the decision unit. On the other hand, if the decision is to be made on each unit or stratum, then each unit or stratum is the decision unit. In this case, representative sample(s) need to be taken from each unit or stratum as if the unit or stratum is the population.

6.4.2.1 If the units or strata are relatively small in size or too numerous to take many samples per unit or stratum, composite sample(s) can be taken from each unit or stratum to increase precision without introducing bias. Alternatively, if precision is not a concern and there is sufficient professional expertise to avoid bias, a judgment sample(s) can be taken from each unit or stratum.

6.4.3 *Heterogeneity*—Heterogeneity is discussed in greater detail in Guide D5956.

6.4.3.1 The degree and extent of population heterogeneity affect potential bias and precision in the samples. Population heterogeneity can be viewed at least in three different ways:

(1) When the population is heterogeneous in a random manner in only the distribution of the concentration, but not in the physical materials such as particle sizes, designs such as a simple random sampling design will generally produce samples with minimal bias. Its precision will then depend on the number of samples taken;

(2) When the population is randomly heterogeneous in concentrations due to large differences in the materials such as particle size, a simple random sampling design may still be effective if the sample volume/weight and sampling equipment are chosen to accommodate the largest particles and thereby prevent introduction of bias; and

(3) If the population is systematically heterogeneous, such as the presence of stratification in concentrations, then a simple random sampling design may not be biased, but will be less precise than an alternative design such as stratified random sampling.

6.4.3.2 Heterogeneity in the population affects the sampling variance. Sampling variance is a function of factors such as the population heterogeneity and the sample volume or weight. It is clear that the more heterogeneous the population is, the larger the inherent sampling variance is. It is also clear that samples of smaller volume or weight will have a higher sampling variance than those with greater volume or weight. However, the reduction in sampling variance due to increased volume or weight may eventually reach a limit. Determination of the optimal sample volume or weight is beyond the scope of this guide.

6.4.3.3 The proper procedure is to first determine the right sample volume or weight, then to determine the number of samples needed for the chosen sample volume or weight.

6.4.3.4 Since stratification as a phenomenon of population heterogeneity is fairly common, it is discussed in greater details as follows.

6.4.4 *Stratification*—There are generally three types of stratification affecting sample representativeness. One is a stratification in the distribution of the contaminant concentration distribution alone. The second is a stratification in sampling materials or matrices alone. The third is a combination of both types. Stratification of any type is not a big problem regarding sample representativeness if each stratum is a decision unit. In that case, the units in a stratum are by definition relatively similar, apart from the random variations in concentrations. A simple random sampling design can be used to obtain representative samples (unbiased) for each stratum. The question of sample representativeness becomes more complicated when a decision is to be made over all the strata in the population.

6.4.4.1 *A Single Representative Sample in a Stratified Population*—When the objective is to obtain a single (physical) representative sample of all the strata, the sample must be a composite of individual samples from the strata (for example, at least one individual sample per stratum). Here the volumes or weights of the individual samples should be proportional to the relative stratum sizes. The composite sample so obtained would be unbiased. However, since there is only one composite sample, precision of the composite sample cannot be estimated. If there are existing data on the precision of the individual samples in the strata, then the precision of the composite sample can be inferred from the precision of the individual samples by theoretical or empirical relationship. See Guide D6051.

6.4.4.2 *A Representative Set of Samples*—When the population is stratified, a set of samples obtained by statistical designs such as stratified random sampling, where the number of samples to be taken from the strata are proportional to the relative sizes of the strata, is unbiased and more precise than a set of samples taken without considering the stratification.

6.4.5 *Parameter(s) of Interest*—This refers to the statistical parameter such as mean or variance of the population. It is often used with a characteristic such as concentration of a constituent(s) of the population. An example is the mean (parameter) concentration (characteristic) of lead (constituent). Another example is a population of mixture of silt-size calcium carbonate particles and large cobble-size particles of calcium carbonate. The interest here could be in the mean (parameter) particle size or chemical composition (characteristic) of calcium carbonate (constituent), depending on the study objective.

6.5 *Develop a Sampling Design*—The objectives of a sampling design are to minimize bias and achieve a desired level of precision. Precision and bias are an issue at various stages of the process of inferring from the samples to the population. The first stage is the act of obtaining the physical samples. The second stage is the act of analyzing the physical samples and translating them into data. The third stage is the use of statistical method to infer from the sample data to the population. At the first stage, the main concerns are sampling precision and bias. At the second stage, the concerns are measurement of precision and bias. At the third stage, the concern is statistical bias.

6.5.1 At the first stage of obtaining physical samples, the issues of precision and bias are sometimes grouped together as sampling design issues.

6.5.2 Bias at this stage is often called the sampling bias. Sampling bias is the systematic difference between the value inherent in the physical samples and the true population value. The word "inherent" is used because at this point the physical samples have not been translated into data.

6.5.3 The phrase "systematic difference" implies a persistent difference in long-term average or expectation, not the occasional random difference. Representative samples, apart from the issue of precision, are obtained when this long-term expected difference is zero or nearly so.

6.5.4 Since the true population value is typically not known, sampling bias cannot be assessed. However, efforts to minimize sampling bias can be attempted in at least two areas:

6.5.4.1 *Proper Statistical Sampling Design*—Statistical sampling design has to do with where and how samples are to be taken, where equal probability of selecting any of the units or items in the population is often a primary requirement. If the probability of selection is not equal, it is highly likely that bias will have been introduced into the physical samples so obtained. Depending on the layout of the population, designs such as simple random sampling or stratified random sampling can be used.

6.5.4.2 *Proper Sampling Procedures and Sampling Equipment*—This includes proper procedures for compositing, subsampling, sample preparation and preservation, and proper use of the chosen sampling equipment. This is a major source affecting precision and bias, especially bias.

6.5.5 In the case of precision, it can be controlled by things such as the number of samples taken, the use of composite samples, or more precise sampling techniques. Often, the number of samples to take is considered the key design issue. Some considerations regarding precision are:

6.5.5.1 If a population is relatively small compared to the sample mass/volume and the distribution of the characteristic of interest is random, it may be appropriate to collect a smaller number of samples by a random or systematic sampling approach, and

6.5.5.2 If a population is relatively large compared to sample mass/volume and the characteristic of interest is not randomly distributed (for example, stratified), a greater number of samples and a stratified sampling approach may be needed.

6.5.6 *Compositing*—Compositing is the combination of two or more individual physical samples into a single sample. It is often used to reduce the analytical costs, while maintaining or increasing precision relative to the individual samples (see Guide D6051). Bias may or may not be introduced in compositing, depending on the study objective and the physical means of compositing. For example:

6.5.6.1 If the study calls for the estimation of the population variance (or standard deviation) of individual samples, then composite samples will surely underestimate the population variance, and

6.5.6.2 If the physical means of compositing changes the characteristics of the samples, then bias may have been introduced (unless such changes are part of the study design).

6.6 *Subsampling*—Sampling bias can be introduced in subsampling unless the same proper sampling protocol is followed as in taking samples from the original population.

6.6.1 After the physical samples have been obtained and before they are measured, bias can be prevented by following proper sample preservation and preparation procedures. It is not important whether these procedures are viewed as part of the sampling process or as part of the measurement process. It is only important in following the proper procedures to prevent bias.

6.7 *Measurement of Precision and Bias:*

6.7.1 The measurement process, like the sampling process, also consists of a random error and a systematic error. The random errors define the degree of measurement precision, and the systematic error defines the degree of measurement bias.

6.7.2 Like sampling precision, measurement precision is controlled by things such as the number of replicate analyses performed per sample and refinements of the analytical method.

6.7.3 Measurement bias is a systematic difference between the sample value produced by the measurement process and the true population value, assuming that the physical samples are unbiased before the analysis. The bias can come from contamination, loss or alteration of the sample materials, systematic errors in the measurement device, or from systematic human errors.

6.7.4 Often the measurement bias can be reasonably estimated in a laboratory testing setting when the true value is known. Laboratory samples spiked with known quantities of a chemical or certified reference standard can often be used to assess potential measurement bias. Minimization or adjustment for such estimable bias in the measurement process is essential in order to obtain data that are unbiased. When estimation of bias is not possible, care in measurement protocol and training is probably the only recourse.

6.7.4.1 It is important to note that, when inferring from the sample data to the population, all the sources of imprecision, including sampling, subsampling, and measurement, need to be combined. The process of accumulating these sources of variation is sometimes called the "propagation of errors." The determination of the optimal numbers of samples, subsamples, and replicates is an issue of optimization and is not covered in this guide.

6.8 *Statistical Bias*—Statistical bias can result from an inappropriate sampling design or inappropriate estimation procedures, or both.

6.8.1 *Selection Bias from Sampling Design*—In the course of taking the sample, if the population units do not have the same probability of being selected, bias can be introduced. This bias can be prevented or minimized when a statistical sampling design is carefully selected, based on the study objective and the layout of the population. Some possible designs are the simple random sampling design and the stratified random sampling design.

6.8.2 *Estimation of Bias from Estimation Procedures*—This bias occurs when the expected value of the statistical estimator does not equal the true value.

6.8.2.1 Estimation bias can occur when the wrong statistical distribution of the data is used. For example, if the normal distribution assumption is used when the true data distribution is lognormal, the interval estimate of the mean concentration will be a biased estimate against the true interval. Thus, the expected value of the estimator will not be equal to the true value. To avoid this potential bias, it is wise to check the data distribution.

6.8.2.2 Estimation bias can also occur when a wrong statistical estimator is used. For example, if the sum of squares of deviations from the sample mean divided by the number of samples (that is, $\sum_{i=1,n} (x_i - \bar{x})^2/n$) is used to estimate the population variance, then this estimator is biased (its mathematical expected value is not equal to the population variance). If its denominator is modified to be $(n - 1)$, then it is an unbiased estimator. For an unbiased statistical estimator, the reader is advised to check with a statistician.

7. Attributes of Representative Samples

7.1 The attributes of a representative (physical) sample or a representative set of (physical) samples can be described in the chronological order in which samples are taken. Note that these attributes apply only to how representative the physical samples are of the population. This corresponds to the upper half of Fig. 1.

7.2 *Design Considerations:*

7.2.1 A well-defined target population. The target population includes all the population units as determined from the stated problem.

7.2.2 The sampled population equals the target population in their spatial or temporal boundaries, or both. The sampled population consists of the population units directly available for measurement.[3]

7.2.2.1 When all the population units in the target population are accessible and directly available for measurement, then the sampled population is identical to the target population in its spatial or temporal boundaries, or both.

7.2.2.2 When not all the population units are directly available for measurement, then the inference from the sample is made to the sampled population, not the target population.

7.2.3 Size (weight or volume) of the sampling unit is well defined.

7.2.3.1 The population can be divided into various sizes (weight or volume) of population units. The size of the sampling unit is the size of the population unit most appropriate for the sampling purposes.

7.2.3.2 The appropriate size of the sample is determined by degree of heterogeneity of the materials to be sampled, such as particle size or shape.

7.3 *Sampling and Measurement Considerations:*

7.3.1 Correct sampling procedures are followed to minimize sampling bias.

7.3.1.1 Absence or minimization of bias is a key attribute of representative samples. Sampling bias can be minimized by following correct sampling procedures. Correct sampling procedures have two components.

(1) A sampling procedure that maximizes the potential of population units having equal probability of selection as sampled, and

(2) Correct sampling procedures. This includes the selection of appropriate equipment and proper use of that equipment.

7.3.2 Sample integrity is maintained during sampling and before chemical analysis.

7.3.3 If subsampling is performed, correct sampling procedures are followed to minimize sampling bias.

7.3.4 Sample preparation errors such as contamination and loss or alteration of constituents are prevented or minimized.

7.3.5 The samples, in the end, collectively reflect the target population within the context of the problem.

7.3.6 These attributes can be summarized into three broad categories:

7.3.6.1 A well-defined population,

7.3.6.2 Correct sampling procedures, and

7.3.6.3 Samples collected in the context of the stated problem.

8. Practical Considerations

8.1 *Sampling Equipment*—The choice of appropriate sampling equipment can be crucial to the task of collecting a representative sample or a representative set of samples.

[3] Gilbert, R. O., *Statistical Methods for Environmental Pollution Monitoring*, Van Nostrand Reinholt Co., New York, NY 1987.