NUREG/CR-1284
PNL-3264
RS

# METHODS OF DETERMINING AND CONTROLLING BIAS IN NUCLEAR MATERIAL ACCOUNTING MEASUREMENTS

R. J. Brouns
J. A. Merrill
F. P. Roberts

June 1980

Prepared for

Safeguards Standards Branch
Office of Standards Development
U. S. Nuclear Regulatory Commission
Washington, D.C. 20555
NRC FIN No. B20149

Pacific Northwest Laboratory
Richland, Washington 99352

8012230323

# ABSTRACT

This report describes the common sources of bias associated with measurements made for nuclear material control and accounting purposes and presents recommended approaches to monitoring and controlling such errors. The effect of measurement error in material balance accounting is discussed to place measurement bias in perspective. The recommended program for monitoring and controlling bias includes calibrations, routine measurements of reference standards, method testing and comparative analyses.

## CONTENTS

## FIGURES

# SUMMARY

This report describes the common sources and the effects of measurement bias in material balance accounting and the general and specific methods for monitoring and controlling bias in measurement processes for special nuclear material (SNM). The study was performed for the Nuclear Regulatory Commission to help improve measurement control programs in the licensed nuclear industry.

Bias of measurement processes should normally be determined and controlled by calibrations and routine measurement of reference standards. As a supplementary bias evaluation procedure, comparative measurements should be made of typical process material samples using an alternative measurement method that is accepted as a standard method and is independently calibrated with reference standards. In some cases, designed experimental tests of the measurement process should be used to evaluate the impact of potential sources of bias and assist in the development of bias-free procedures. The accuracy of sampling is generally established in this manner.

While bias is expected to be eliminated by calibration, within the limitations imposed by the random error of the calibration procedure, every measurement process should be continually tested to monitor the validity of the calibration under current conditions and to determine bias correction factors if required. The frequency of bias tests should be established for each type of measurement system on the basis of historical monitoring data. The potential impact of an uncorrected measurement method bias on the inventory difference (ID) of the facility material balance should also be considered. Where the impact could be appreciable, monitoring for bias should be more frequent than otherwise. Control charts should be used to monitor the stability of the calibration.

To help avoid sources of error stemming from composition differences, the reference standards used for routine bias tests should be similar in form and composition to the process samples for which the measurement process is used. In addition, the measurements of reference standards should be made using the same procedures, equipment, instruments, locations and personnel as for routine

process measurements. The measurement results must be traceable through an unbroken chain of comparisons to national standards or national measurement systems.

As an aid in choosing or developing suitable calibration methods and experimental tests for each measurement process, the user should be very familar with the chemical and physical bases of each measurement procedure and the known sources of bias in the use of the procedure. This report reviews the potential sources of bias in nuclear material mass measurements, sampling, chemical and nondestructive assays and isotopic ratio measurements. In addition, a list of references is included that provides supporting information and procedural details for most of the measurement systems used in the nuclear process industry at the present time.

# 1.0 INTRODUCTION

Close control of special nuclear material (SNM) measurement quality is necessary to ensure that a loss, theft or diversion of SNM will not be masked by errors in the measurement data used in the periodic material balances that are performed to test SNM control and accounting. The detection and minimization of bias in the measurements is particularly important to avoid a biased inventory difference (ID) in the SNM material balance. Even a small relative bias in a measurement process may amount to an appreciable bias in terms of quantity of SNM when that measurement process affects large quantities of SNM over a period of time. Since a positive ID that is statistically significant indicates a material loss, biases that increase ID will increase the chances of a false indication of loss. On the other hand, biases that decrease ID may mask and prevent detection of actual losses, thefts or diversions.

The importance of controlling biases is generally accepted and the use of calibrations to insure accuracy of measurements is widely practiced. Nevertheless, because of the potential for variable biases in some measurement processes and the importance of even relatively small biases in material balance data, there is a need for continuing vigilance in testing and recalibrations.

Federal regulations applicable to licensees who possess certain quantities of special nuclear material(a) require licensees to establish and maintain a program "for the continuing determination and control of the systematic and random errors of measurement processes...".(2) This program must provide for the "ongoing use of standards for calibration and control of all applicable measurement systems".(1) Some guidelines and recommended practices directed toward these requirements have been published previously.(3,4,5) This report reviews the sources of bias in typical nuclear material control and accounting measurements and presents suggested methods for monitoring and controlling bias. These methods include recommendations of bias monitoring techniques such as

---

(a) A quantity of material exceeding one effective kilogram of SNM for use in activities other than those involved in operation of a nuclear reactor licensed pursuant to 10 CFR Part 50, those involved in a waste disposal operation, or as sealed sources.(1)

periodic recalibrations, frequent measurements of reference standards, special tests for conditions that may lead to bias, and periodic comparative measurements by other methods of known accuracy.

# 2.0 THE NATURE OF MEASUREMENT ERROR

The following discussion is presented to define the basic concepts of measurement error, particularly the systematic error or bias component.[a] It is recognized that although the concepts are generally accepted, there is considerable variation in terminology and in the statistical handling of systematic errors. The purpose here is to place the concepts and terminology in a perspective relevant to the remainder of this report.

Whenever a chemical or physical attribute (length, weight, volume, chemical composition, etc.) for a material or item is measured, the value obtained will be in error because of random variations in the measurement process and perhaps because of an offset relative to the correct value of the attribute. Inherent limitations of measurement processes cause such error. In a simple case, the error for any particular measurement is composed of two components: the systematic error component and the random error component. The two components are different in that the random error component varies from measurement to measurement while the systematic component remains fixed for all measurements in the data set. To clarify the distinction between these components, the following model for a simple case will be useful.

Let

$x_i = \mu + \beta + \varepsilon_i$ for $i = 1, 2, \ldots, n$ (n measurements taken on an individual item) (1)

where

$x_i$ = the $i^{th}$ measured value

$\mu$ = the true value for the attribute being measured[b]

---

[a] In everyday usage, the systematic error component is frequently also referred to as the bias. The term bias is defined in 10 CFR Part 70, Section 70.57 (a)(6), as a systematic error <u>whose value has been determined</u> from the deviation of the mean of a measurement process from a reference value, and <u>whose effect can be corrected for</u> (emphasis added).

[b] The traditional expression "true value" is used here because of its general familiarity. Some writers prefer "target value," or "sought after value".(6)

$\beta$ = the systematic error component or bias

$\varepsilon_i$ = the $i^{th}$ random error component (frequently the $\varepsilon_i$ are assumed to be independent and distributed normal ($o$, $\sigma_\varepsilon^2$ )).

Subtracting $\mu$ from both sides of the equation, the right side becomes $\beta + \varepsilon_i$, which is equal to the measurement error. The systematic error component is represented by $\beta$ and, in this simple model, is a constant. The random error component is represented by $\varepsilon_i$. This term is not constant but may vary in both magnitude and sign from measurement to measurement.

If several measurements are made on the same sample or item, the results generally will show a variation. As the number of measurements increases, the average measurement value will approach a constant or "limiting" value. The limiting value of the average may not coincide exactly with the true value and this difference between the limiting value and the true value is the bias of the measurement process.

If the bias were known exactly, each measured value could be corrected. For such a hypothetical case, the measurement error would be reduced to just the random component. In actual practice, however, the bias is never known exactly but must be estimated using a process which has error.

A typical method for estimating the bias is to use the measurement process to make a number of measurements on a reference standard. The difference between the average of these reference standard measurements and the assigned value for the standard is an estimate of the bias. (The actual or true bias may be defined as the difference between the expected value of a random variable and its corresponding correct value.) In the uncomplicated cases to be discussed first, the closeness of this estimate to the true value of the bias of the measurement process when applied to materials measured routinely depends on: 1) the size of the random error standard deviation for the measurement process, 2) the number of measurements made on the standard, 3) how closely the reference value assigned to the standard corresponds to the true value for the standard, and 4) how closely the standard resembles in composition, form, or chemical interferences the items or materials to be measured routinely.

It must be recognized that estimates of the bias are subject to uncertainty and that any estimate is merely one from a distribution of all possible estimates. From this it follows that there is a variance associated with the estimate. Such variances need to be included in an appropriate way when combining error sources to obtain the limit of error of an inventory difference (LEID).

A simplified example follows to show the sources of uncertainty in bias estimates and how these affect the uncertainty associated with a measurement process. Suppose that during a given period a number of plant samples are taken from the process and measured for a chemical property. Let the model for the plant sample measurements be:

$$x_{ij} = \mu_i + \beta + \varepsilon_{ij} \tag{2}$$

where

$x_{ij}$ = the value for the $j^{th}$ measurement of the $i^{th}$ process sample

$\mu_i$ = the true value for the $i^{th}$ process sample

$\beta$ = the bias of the measurement process

$\varepsilon_{ij}$ = the random error of the $j^{th}$ measurement of the $i^{th}$ sample and $\varepsilon_{ij}$ has a mean of zero and a variance of $\sigma_\varepsilon^2$

Suppose further that a standard is measured n times during the period in order to estimate the bias. Let the model for the measurement of the standard be:

$$y_i = T + \beta + \eta_i \tag{3}$$

where

$y_i$ = the $i^{th}$ measured value of the standard

$T$ = the true value of the standard

$\beta$ = the bias component (i.e., the same value as in Equation 2)

$\eta_i$ = the random error for the $i^{th}$ measurement of the standard where $\eta_i$ has a mean of zero and a variance of $\sigma_\eta^2$

Also, let the assigned reference value of the standard be R, which is not necessarily equal to T.

The random error components $\varepsilon$ and $\eta$ are generally of the same magnitude but are given different symbols to permit a distinction to be made between deviations encountered on plant samples and on standards. The bias, $\beta$, is assumed to be the same for measurements of samples and standards, the standards being chosen from a range of standards to closely approximate the sample composition.

An estimate of the bias, $\hat{\beta}$, is obtained by subtracting the assigned reference value, R, of the standard from the average, $\bar{y}$, of the results from the standards measurements, i.e.,

$$\hat{\beta} = \bar{y} - R \tag{4}$$

The value for R is not exactly equal to T, the true value of the standard, because the measurement process used to arrive at the assigned value has random errors and bias associated with it. Therefore, T differs from R by some value $\Delta$, i.e.,

$$\Delta = R - T \tag{5}$$

The value of $\Delta$ is, of course, not known but knowledge concerning the potential magnitude of $\Delta$ in the assignment of reference values is needed, since from Equation 4,

$$\text{Var}\ (\hat{\beta}) = \text{Var}\ (\bar{y}) + \text{Var}\ (R) = \frac{\sigma_\eta^2}{n} + \text{Var}\ (R) \tag{6}$$

Thus, the variance of $\hat{\beta}$ stems from both the variance of the value R assigned to the standard and the variance of $\bar{y}$.

A bias-corrected plant measurement is $x_{ij} - \hat{\beta}$, or using Equation 2,

$$x_{ij} - \hat{\beta} = \mu_i + \beta + \varepsilon_{ij} - \hat{\beta} \tag{7}$$

Thus

$$\text{Var}\ (x_{ij} - \hat{\beta}) = \sigma_\varepsilon^2 + \frac{\sigma_\eta^2}{n} + \text{Var}\ (R) \tag{8}$$

Therefore, the sources of variation of a bias-corrected measurement are the variance of the measurement process, $\sigma_\varepsilon^2$, and the variance associated with the bias correction. The variance of an average of m bias-corrected measurements on a sample is

$$\text{Var}\ (\bar{x} - \hat{\beta}) = \frac{\sigma_\varepsilon^2}{m} + \frac{\sigma_\eta^2}{n} + \text{Var}\ (R) \tag{9}$$

where

m = the number of measurements made on the sample

n = the number of measurements made on the standard

Thus, the variance of an average of bias-corrected measurements can be reduced by increasing the number of measurements on the sample, by improving the measurement method to reduce $\sigma_\varepsilon^2$, and by reducing the variance of the estimated bias. The variance of the estimated bias can be reduced by increasing the number of measurements made on the standard, by improving the measurement method to reduce $\sigma_\eta^2$, and by using a standard with a smaller variance of R. Frequently, however, the limit or bound given for the reference standard is sufficiently small that it has no appreciable effect on the uncertainty of the bias estimate.(a)

Bias estimates of very well-characterized measurement processes will include all sources of systematic error. In many practical cases, this is not attained and the error model is usually more complex than illustrated above. Bias can occur if the measurements used for the calibration of the method or the bias estimates are made under conditions not identical to those encountered when measuring process materials. Differences in procedural steps, environmental conditions, or such factors as apparatus and operators can introduce offsets in the measurement process. Measurements also can be biased because the standards do not respond to the measurement process in exactly the same way as process materials due to differences in composition or other material characteristics. Such biases could be estimated by careful characterization of the measurement process using such techniques as methods testing and comparative measurements (see Sections 4.2 and 4.3). The measurement process often can be modified to eliminate the cause of bias or to reduce its effect to a negligible level. In

---

(a) The limits assigned to R, the reference value of a standard are normally given by the one who prepared and calibrated the standard. They are derived from the variance of the calibration process (the measurement data), the limits or bounds of the standard(s) of higher hierarchy used, and estimates of the uncertainty introduced by all other possible sources of error in the calibration process.

those cases where this is not practical, it may be possible to estimate limits (or bounds) for the bias for each source that has a non-neglible effect on measurement results. These limits are to be included in the overall uncertainty of the measurement process.(a)

---

(a) If the effect of a source of bias can be quantified (i.e., the direction and magnitude of the resultant error can be determined), the measurement result can be corrected.

# 3.0 SOURCES OF BIAS

## 3.1 VOLUME MEASUREMENTS

Volume measurements for SNM accounting are normally made in calibrated tanks equipped with instruments for determining liquid level. Volume is calculated from liquid level readings using a calibration curve or equation. The calibration equation of a linear volume measurement tank is:

$$h_L = a + bV$$

The working equation for volume measurement is:

$$V = (h_L - a)/b$$

where

$V$ = volume of liquid

$h_L$ = liquid height in the tank

$a$ & $b$ = calibration constants, namely, the intercept and slope of the calibration curve, respectively(a)

The liquid level, $h_L$, may be measured by any of several instrument systems. The most common system for remote tanks in high-radiation zones is the pneumatic probe system, in which the primary measurements are liquid head, measured by a differential pressure instrument or simple liquid-filled manometer, and density. The mathematical function for a manometer readout is:

$$h_L = h_M \rho_M / \rho_L$$

where

$h_M$ = the liquid height in the manometer

$\rho_M$ & $\rho_L$ = the densities of the liquids in the manometer and the tank, respectively(b)

---

(a) Volume measurement tanks that are not strictly linear may require curvilinear calibration equations or a set of two or more linear equations to obtain the best representation of liquid height to volume.

(b) If the liquid in the tank is homogeneous, the density can be obtained from the pressure difference between two levels below the surface of known difference in elevation.

Other common liquid level measurement methods are sight gage, float and ultrasonic ranging.(7,8) Most of the sources of volume measurement error are the same for any liquid level measurement method.

Tank calibrations are expected to be stable for long periods of time if the tank and its internal equipment are sufficiently rigid. However, the calibration curve of a tank may change because of a change in the dimensions of the tank from relief of stress or physical distortion under load. In addition, the calibration curve can be altered by corrosion of the lower pressure sensing tube and other internal surfaces of the tank and by replacement of heating or cooling coils or stirrers with units of different dimensions. The total liquid capacity of the tank can also change due to buildup of sediment. This usually only affects the heel volume (the residual contents of the tank not normally removable by pumping), which is part of the tank inventory but which should not affect the accuracy of volume transfers.

Any of the instruments used in the measurements of liquid level, such as pressure sensors and thermometers, can give biased results. Some possible causes of error are electrical disturbances, degraded or failed instrument components, or uncompensated changes in ambient temperature. Liquid-filled manometers, widely used to measure the pressure differentials, depend for their accuracy on an accurate knowledge of the density of the manometer fluid.(a) The density of manometer fluids can be affected by contamination, aging and radiation damage. These densities also vary with temperature, but adjustments for such variation can be made routinely. Because of these potential sources of error, manometer fluid is replaced or the density remeasured frequently.

Lack of care and skill by operators can also be a source of volume measurement error. Some operators may make instrument adjustments or read instruments in a manner significantly different from others. Such reading differences do not necessarily cancel each other, even in measurement of volume transferred

---

(a) The solution level in the tank is calculated from the static pressure differential and the ratio of the manometer fluid density to the tank solution density. The measurement of the latter also may be based on a manometer fluid density value.

which involves the difference between two liquid level readings, because these two readings may often be made by different operators. Consequently, operator skill and adherence to approved procedures are important.

Process equipment other than measurement tanks should be calibrated for measurement of liquid holdup volumes, unless all holdup is to be flushed to measurement tanks for inventories.(9) Holdup measurements can be biased, however, despite thorough calibrations or transfers to measurement tanks, because of such sources of error as incomplete draindown of liquid in the vessels and transfer lines, and changes in the capacity of calibrated vessels caused by accumulations of sediment.

## 3.2 MASS MEASUREMENTS

Potential sources of bias in weighings are:

- environmental conditions such as temperature, humidity, vibration and dust that are different from conditions that existed during calibration
- a calibration change due to corrosion or damage to the scale mechanism
- systematic operator error.

These sources of error are applicable to both gross and tare weighing processes. A bias in the weighing process may affect both weighings equally, in the absolute sense, in which case the error will cancel for net weights, but such cancellation cannot be assumed.

The sensitivity of scales to environmental influences can be determined and precautionary measures instituted. Significant calibration shifts can be detected through a monitoring program with working standards (see Section 4.1 and Section 5). Bias can be caused by operator differences in following procedures for cleaning and drying the objects to be weighed and in reading graduated scales or estimating null points.(a)

---

(a) Operators may also make mistakes in weighing, recording data and calculating results, but these are more likely to be random than systematic. The use of automatic printing or digital readout scale will reduce the frequency of reading errors and mistakes in recording.

A common practice in weighing is to round the observed weight to fewer digits than actually observed. This practice may introduce an additional component to the random error variance in weighing but usually will not cause a bias.(a) However, if calibration data are rounded, the added error will affect the standard deviation of the calibration value, which will affect, in turn, the standard deviation of all process item measurements based on that calibration.

Excessive rounding can cause a bias in certain circumstances. For example, consider the case of a group of items whose average mass is 100.4 g and for which all or nearly all observed values fall between 100.0g and 100.5g. In this case, rounding of either of the observed values or the mean to the nearest gram would result in a 0.4 gram bias.

## 3.3 SAMPLING

Potential sources of bias in sampling are:

1. process and material effects such as:
   - variability or inhomogeneity of the material sampled
   - segregating behavior of powder mixtures (e.g., the sampling procedure may tend to select material preferentially by particle size fraction or density)
   - change in composition of samples while awaiting analysis
   - the presence of solids in solution samples
2. procedural effects such as:
   - inadequate protection of samples from air, moisture and airborne contaminants during sampling and while awaiting analysis
   - contamination of samples from inadequately cleaned or dried equipment and sample containers
   - incorrect sampling plan or procedure

---

(a) Guidelines on rounding practices are available in Reference 10.

3. improper designs or malfunctions of sampling and mixing equipment
4. failure of operators to follow prescribed sampling and sample protection procedures.

These sources of error may affect not only the primary sampling operations but subsequent sample handling operations as well, such as preparing, blending, and subdividing the gross sample, and preparing the laboratory sample.

Variability of the composition of material within a batch or lot would not necessarily result in biased samples. If the variability were random, the sampling error would vary in magnitude and direction from sample to sample but on the average the results would not be biased; if a sample is a composite of many subsamples from the lot, the results would approach the lot mean as the number of sample increments is increased. However, if the composition varied systematically, such as from place to place in containers or from time to time in the production flow, samples could be biased if random sample selection procedures were not used.(a) Consequently, bias can be the result of improper sampling procedures or the failure of operators to follow the sampling plan.

Process materials cannot always be blended thoroughly before sampling. Processing or equipment limitations or criticality constraints on batch sizes may preclude thorough mixing. Blending problems also occur with powder mixtures that tend to segregate according to particle size or density,(11,12) with multiphase systems, and with heterogeneous materials such as scrap, waste, rags, paper and plastic, contaminated air filters, and ion exchange resins. In such cases, the formation of composites from subsamples obtained using stratified random sampling will be helpful.(4)

Composition changes in a sample awaiting analysis can result from such effects as evaporation (liquid or two-phase samples), hydration or dehydration, oxidation or deoxidation, precipitation, and the formation of polymers (plutonium). Sample evaporation can occur during the sampling operation as well, and vacuum-lift sampling devices can greatly increase such evaporation.

---

(a) That is, sampled in such a way that every portion of the batch or lot had the same probability of being sampled. (For additional details see Reference 4.)

Equipment for remote liquid sampling often includes considerable piping, flush lines, steam or air jets, and retainer cups or wells that can be a source of sample contamination or dilution. If the system is not flushed adequately with process solution before samples are drawn, dilution or cross contamination of samples is likely.

## 3.4 LABORATORY ANALYSES

Chemical and physical analyses of materials may involve many steps and manipulations, each having the potential for introducing error into the overall measurement. Common steps are drying, mixing, subsampling, weighing, dissolving, separating, purifying, and assaying. Potential sources of bias in analyses are:

- error in calibrations of the measurement process or instruments, reagents, and equipment such as balances and volumetric ware
- the presence of interfering constituents in samples
- differences in composition or form between standards and process samples that affect method accuracy
- incomplete dissolution of samples
- uncompensated changes in composition of standards and reagents, such as by evaporation or contamination
- failure of operators to follow prescribed procedures.

Chemical interference is a common and troublesome source of error. Either the presence of the interference or its effect may not be known. Analytical method development and testing normally include a careful study of the effects of potential variations in material composition and impurity levels on the accuracy of the method. Method modifications are made that eliminate the interferences if possible, or correction factors are determined. The presence of sporadically occurring impurities in process samples that cause analytical interferences can be monitored by performing a general analysis (e.g., emission spectrography). When such an analysis shows that an interfering substance is present, either a bias correction can be applied, if it is known, or a countermeasure such as sample pretreatment and separation can be carried out.

Incomplete sample dissolutions may result from inadequate procedures, operator error, or differences in solubility behavior between different lots of material. For example, the resistance of plutonium dioxide to dissolution is affected by calcining temperature and time.(13)

## 3.5 NONDESTRUCTIVE ASSAYS (NDA)

Nondestructive assays of the SNM content of nuclear materials are achieved by measuring the radiations or the heat generated by the radioactive decay of the isotopes of uranium or plutonium, or their decay products, or some combination of these. The radiation may be spontaneously occurring or may be induced by gamma or neutron irradiation. Methods based on the former are called passive assay techniques, and those based on the latter are called active assay techniques. Descriptions of the designs, performance and applicability of existing NDA instruments are available in several publications (see, for example, References 14-16).

All currently used NDA systems respond to a variety of factors in addition to the quantity of SNM in an item that may result in bias. The principal factors can be categorized as:

1. instrumental and environmental effects coupled with calibration procedures that do not provide for control or compensation for these effects
2. differences between the composition of measured items and the composition of the calibration standards.[a]

Some of the instrumental and environmental factors which may cause bias are changes in the:

- intensity of a radiation source or in its energy spectrum
- instrumental response, due to aging components or to varying ambient conditions such as temperature, humidity or power
- background radiation level (either from internal or external background sources)

---

[a] This is one of the principal causes of item-to-item error (see References 16 and 17).

- the position of a detector(s) or a radiation source relative to an item measured.

As an example of an effect due to ambient conditions, consider the case of using, without ambient temperature control, an instrument whose response is affected by temperature. The instrument will give biased results when the mean temperature of the room in which the instrument is used is not maintained at the mean temperature at which it was calibrated.

Potential sources of bias in NDA measurements due to differences between measured items and calibration standards include differences in:

- SNM density or distribution
- the form, composition, density or distribution of extraneous materials that affect the attenuation of radiation, or the amount of extraneous radiation emitted
- the isotopic composition of the SNM
- the container size, geometry, wall thickness, or composition.

The effect of differences in form, composition or density between items measured and calibration standards is generally difficult to monitor and control. The composition differences may not affect the reproducibility of repeated measurements of a given item, but may cause consistent differences in response from item to item that are not the result of a difference in SNM content.

As an example of a situation in which such errors can occur, consider an instrument that measures the $^{235}U$ content of containers of oxide powder scrap of varying bulk density. If the instrument response is affected by the density of the material in the containers, the response per unit of $^{235}U$ will vary as the density varies. Therefore, a bias will occur in the assay of a batch of items if the mean density of the material in the batch is significantly different from the mean density of the calibration standards.

Similar effects can be caused by such characteristics of items as changes in the kinds of extraneous materials in the items and the distribution of the

SNM and other materials. For each NDA application, efforts are made to fabricate standards having the expected average characteristics (form, composition and distribution) of the items to be assayed. If the chosen characteristics are right, the average result of the measurements of a group of unknown items should be substantially unbiased, although considerable variability from item to item may occur in the assays.

The effects of material variability are expected to be minimal for uniform materials such as fuel pins and pure compounds because the variations in SNM purity and distribution, item geometry and matrix composition (including impurities) are usually small. Less pure and less uniform materials such as scrap and waste can vary in item characteristics over a wide range and error from item-to-item, and possibly the bias of the mean of many items, can be difficult to evaluate.

The main efforts in controlling the accuracy of NDA are directed toward: 1) developing NDA instruments and techniques that are less sensitive to varying item characteristics and 2) characterizing the form and composition of the unknown items so that reference standards of that form and composition can be developed. This characterization is done principally by method testing and analytical methods. Further discussion of this subject is in Sections 4.2 and 4.3 and greater detail may be found in various other publications and guides (See References 14-19).

# 4.0 METHODS FOR DETERMINING BIAS

Usually the bias of a measurement method is determined and controlled by calibrations and routine measurements of reference standards that can be traced to a national measurement system. However, a calibrations and standards program may be supplemented by methods testing and comparative measurements. Methods testing usually consists of a systematic experimental evaluation of the response of a measurement process to variations in the procedure, the environmental parameters and the sample material composition. Comparative measurement is the use of an alternative method of known accuracy to determine the "true" value or composition of sample materials in a paired comparison study. These methods are particularly important for evaluating the accuracy of methods for which suitable standards are not available. This situation is common to chemical analysis in industry and research. However, nationally accepted reference standards are available for the measurement methods important in SNM accountability. Therefore, control of accuracy entirely through calibrations and reference standards is generally possible.

In the area of SNM accounting measurements, method testing and comparative analysis are used mainly for establishing unbiased sampling procedures, testing for bias in NDA and performing tests for diagnostic purposes. Conventional calibration techniques are not applicable to sampling method evaluation and, in NDA, calibration requires verification when one cannot be certain that the matrix and the distribution of the constituents in the reference standards is representative of the process samples.

## 4.1 CALIBRATIONS

### 4.1.1 Procedures

Measurement processes applied to SNM measurements for materials accounting purposes should be calibrated using physical standards whose reference values can be related back through an unbroken chain of comparisons to national standards or nationally accepted measurement systems.(5) Calibrations normally result in bias correction factors or calibration lines or curves which are used to relate the measurement process output to the value of a characteristic of the material. The processes of calibration, standardization, recalibration and determination of bias corrections or bias correction factors are basically identical, although initial calibrations and subsequent recalibrations may differ in such details as the range covered and the extent of testing for stability and linearity. For simplicity, the term "calibration" will be used to refer to all of these procedures.

Calibration procedures applicable to SNM measurements are found in the guides and standards of the American Society for Testing and Materials, the American National Standards Institute, the U.S. National Bureau of Standards, and the U.S. Nuclear Regulatory Commission, as well as the literature of manufacturers of measurement equipment. Procedures for calibrating scales, balances, and mass standards (weights) are described in References 20, 21 and 22. Volume calibration procedures for tanks are available in References 23, 24 and 25 and volume calibration procedures for laboratory glassware are found in Reference 26 as well as many analytical chemistry textbooks. Mass spectrometer calibration procedures are given in References 13, 27 and 28. Recent summaries of the chemical analysis methods for uranium and plutonium are given in References 13 and 29-31. Guidelines for calibrating NDA instruments are given in References 14, 16 and 19.

Calibrations should be made to an accuracy that is appropriate and adequate for the measurement application. This involves consideration of the attainable precision of the measurement, the quantity of SNM represented by a single measurement, and the total amount of SNM being measured with the calibrated system. Usually, the measurements of standards(a) in calibrating a system are repeated enough times to reduce the estimated standard deviations of the calibration parameters (mean calibration factor, bias correction factor, or slope and intercept of a linear calibration) to levels substantially less than those obtained in measurements of unknowns. Independent replication, preferably distributed throughout the period that the system is used for process measurements, is necessary. Independent replication reduces the standard deviation of a mean

---

(a) Standards used in calibrations will be discussed more fully in Section 4.1.2.

inversely to the square root of the number of measurements ($s_{\bar{x}} = \frac{s_x}{\sqrt{n}}$). However, after about 15 replications have been made, the gain for each additional measurement is small (see Figure 1).

It cannot be established for the general case just how small the standard deviation associated with calibrations should be but it can be determined for each application of a measurement system. This should be done by projecting the effect of each measurement process calibration standard deviation on the standard deviation of the inventory difference.

An appropriate recalibration frequency should be established for each measurement system. The frequency inevitably varies from system to system and may be a function of the type of equipment involved, the accuracy requirements placed on the system, the environmental conditions, and the performance and calibration history of the system. One common practice is to run standards or recalibrate every day for some measurement methods (e.g., isotopic analyses and uranium and plutonium analysis procedures). Regardless of the routine frequency, a new calibration should be performed whenever:

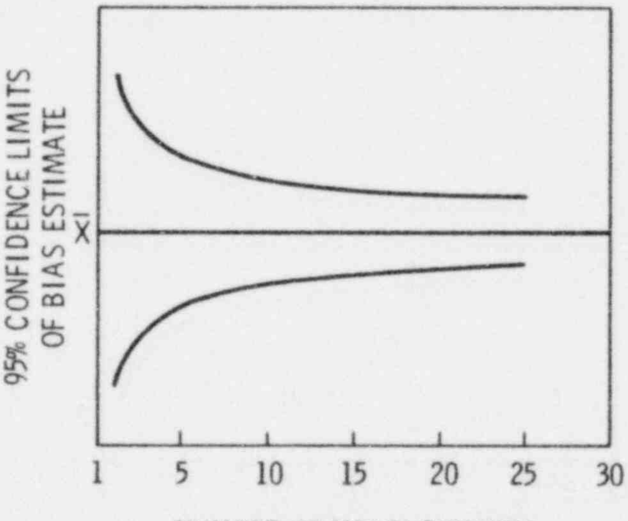

FIGURE 1. Confidence Region of a Bias Estimate ($\sigma$ known) as a Function of n

- repairs or modifications of equipment or changes in procedures have been made
- monitoring data indicate that a significant change in a calibration has occurred
- additional calibration data are needed to reduce the standard deviation of the calibration.

The guidelines given in ASTM Standard D-2865-71 also may be helpful in setting calibration frequencies.(32)

For some measurement systems, calibrations cannot be repeated frequently. An example is a liquid volume measurement tank containing radioactive material or located in a high radiation zone. Such systems may prove to be very stable, but periodic monitoring of calibrations is necessary to demonstrate stability and to guard against unexpected changes (see Section 3.1). Periodic tests for bias of volume measurement systems can be facilitated by having "test measures" or "provers" in place for delivering known volumes of liquid to the measurement tanks. Calibration tests may also be made by 1) comparing volume measurements between interconnected calibrated tanks, and 2) measuring dilution factors for a constituent of the solution or an added tracer.(a) Measurement tanks and instrumentation in isolated locations should be:

- designed and constructed to be very stable with respect to calibration
- carefully calibrated and tested initially to achieve an acceptably small calibration standard deviation
- periodically rechecked to verify the calibration accuracy
- recalibrated at times of plant shutdown and cleanout.

The volume measuring instrumentation can also be checked for bias by making routine direct comparisons with other independently calibrated reference instruments. On-line NDA systems and other instruments that are isolated during

---

(a) Such tests have limited accuracy compared to full scale, multiple-pass tank calibration but they would serve to detect major calibration discrepancies.

process campaigns should be monitored in a similar manner. Experimental and historical data reflecting the stability of a system can provide guidance about frequencies.

Supplementary monitoring techniques that would improve control of accuracy should be used, such as routine measurements of blanks and backgrounds for chemical assays and NDA. Whenever monitoring data indicate a significant departure from normal performance or loss of statistical control (see Section 5), the measurement system involved should be taken out of service until it has been tested, repaired and recalibrated.

Calibration procedures, schedules and results should be documented. The procedures and schedules should have the formal approval of both the personnel responsible for the performance of the measurement systems and of the material control and accounting system.

Calibrations should be performed by qualified personnel. All components of a measurement system (pretreatment, subsampling, detectors, transducers and other instrumentation) should be calibrated as a unit, if possible, rather than separately as subsys[illegible]nalytical systems should be independently calibrated for each type of material to be analyzed except where pretreatments of the material to be analyzed reduce each type of material to the same state and these pretreatments can be shown to introduce no bias.

To avoid possible bias due to compositional differences, working standards usually are made similar in composition and form to the process material they are to represent. Calibrations and bias tests should be made under the same conditions that exist when routine measurements are made. When it is impractical to calibrate under the same conditions as those encountered during routine measurements, sufficient method testing should be performed to determine the magnitude of any bias introduced by the deviation from "normal" conditions. Tests also should be made periodically for between-operator and between-instrument effects and other sources of error that may contribute significantly to the measurement bias by performing designed experiments and statistical analyses such as an analysis of variance.

### 4.1.2 Reference Standards

Reference standards of mass and volume often can be purchased from commercial suppliers. These suppliers can provide certificates of traceability to national standards and the tolerances for the calibrations. In addition, some suppliers and the National Bureau of Standards can provide calibration standards and recalibration services for standards.(22,33,34) Other reference standards and their availability or preparation are described in References 13, 29-32 and 35-37.

"Working" reference standards or working reference materials (WRM) (also commonly called bench, test or check standards) are generally used for day-to-day monitoring of measurement process and equipment accuracy. Such standards should resemble in composition and form the typical process material samples or items to be measured and should be calibrated against primary reference standards. Working standards for chemical and isotopic analysis have been termed working calibration and test materials (WCTM).(35-37)

A variety of WCTM or WRM may be needed to cover the ranges of composition and form that are important for monitoring the calibrations of the measurement systems being used. Since a perfect composition match is not always possible, method tests and comparative analyses are used as supplementary performance monitoring measures, as discussed in Section 4.2 and 4.3. When a range of measurements is involved, calibration standards should be chosen to span the entire working range. Extrapolation of calibration curves is not a recommended procedure.

Mass, volume and NDA working standards are usually very stable and recalibration of them is necessary only occasionally as a precaution against damage, deterioration or tampering. The WCTM for chemical assay are often standardized solutions or powder mixtures that are not stable in all environments. Such standards should be recalibrated relatively frequently. Experimentally determined stability information will help establish the necessary recalibration frequency and suitable conditions of storage. In addition, some guidance on recalibration frequencies is available in published literature.(20,23,29-32,35-37)

Calibration standards should be protected from possible damage, deterioration, misuse or tampering. Locked storage, assigned custodial responsibility, and inventory control are some of the protective measures that are useful. Working standards will, of course, receive the greatest exposure; accordingly, recalibrations should be more frequent than for stored reference standards. A suitable laboratory should be available for preparing and storing reference and working standards and calibration equipment. The laboratory may also be used to perform instrument, equipment and WCTM calibrations. In that case, it should have a well-controlled temperature, humidity, cleanliness, and electrical line stability and be free from excessive vibration.

All working reference standards should be uniquely identified and short-lived standards such as standardized solutions should have expiration dates on the labels. A formal record system for the preparation, calibration and recalibration information on reference standards should be maintained. The approved preparation and calibration procedures for working standards should also be documented.

## 4.2 METHOD TESTING

Method testing is the process of experimentally evaluating the effects on the behavior of a measurement system response of variations in the procedure used, the material or item characteristics such as composition, and the environmental conditions. This is an important technique for identifying the sources of measurement errors and developing ways to reduce error, such as sampling error and bias due to assay method interferences that occur as a result of variations in material composition (see Sections 3.4 and 3.5).

The method testing approach generally consists of systematically studying the response of a method to the range of various conditions that may be encountered in routine application. Statistical experimental design techniques provide efficient and systematic approaches to method testing.(38,39) In particular, factorial experimental designs have proven to be appropriate in such studies. In many cases, the sources of error that are identified can be eliminated simply by controlling or limiting the range of conditions under which the method is used, or by modifying the method to nullify the effect of an interference. Alternatively, the effect of the interference may be quantifiable and the measurement results corrected by determining the amount or intensity of the interference and then applying an appropriate correction. The uncertainty associated with the corrections can be estimated from the methods testing data.

### 4.2.1 Testing NDA Systems

Sample pretreatment for eliminating interfering substances, as used in chemical assay, is not applicable to NDA. Therefore, emphasis is placed on 1) developing NDA methods that are specific for SNM isotopes and are relatively insensitive to the matrix composition and spatial distribution, 2) classifying sample items to restrict matrix compositions for each assay system within narrow ranges, and 3) determining correction factors for interferences. For example, if pellet density affects the NDA response to fuel pellets or pins in a known manner, a correction factor can be applied, provided that the fuel density is measured.

When bias corrections for interferences in NDA cannot be determined, the uncertainty contributed by the interferences may be estimated and reflected in the uncertainty of the results. One method for estimating this uncertainty is to:

1. Identify, a priori, all significant sources of interference in the NDA measurement.
2. Experimentally determine the distribution of the interfering parameters that actually occur in the population of items measured, and determine the empirical relationship between the magnitude of each interference and the effect on the response of the method.
3. Calculate the variance associated with individual item measurements from the information gathered by (2) (i.e., the distribution and magnitude of interference of each variable parameter).

4. Periodically redetermine the distribution of the interfering parameters in items to ensure that the calibration standards are still representative and the distribution functions of the interfering parameters have not changed.

The above approach is applicable only to relatively well-characterized materials that can be expected to have the same distribution of composition over an extended period of time. Fuel rods, pins, pellets, powder, and pure compounds of SNM may be such materials.

Another approach, which may be termed a maximum credible extremes method can be applied.(18) It consists, basically, of the following steps:

1. Identify, a priori, all significant sources of measurement error due to interferences.
2. Estimate the maximum range for each source of interference in the items to be measured.
3. Fabricate standards representing the maximum variability in the characteristics that are sources of interference, as well as standards representing "average" or "typical" items.
4. Measure these standards and prepare a family of calibration curves.
5. Estimate the maximum credible error from the calibration data. (This would include both within- and between-calibration variations.)

This approach is sometimes used to test whether the NDA measurement error may make a significant contribution to the limit error of the inventory difference (LEID). Since the maximum credible error derived with this method does not have a statistical basis, it may be too extreme for inclusion in the LEID calculation. However, if the quantities of SNM are small, such as in wastes, the maximum credible error may be insignificant with respect to the LEID and, therefore, a more refined evaluation of limits of error may not be justified.

### 4.2.2 Sampling Tests

There are three general approaches to using experimental tests of mixing and sampling to determine bias and the standard deviation of the bias estimate. These are:

- comparing measurement data from samples taken by the method under test with data from samples taken by an independent and very reliable sampling procedure. For free-flowing liquids and solids, a composite of several large thief or dip samples from a vessel can be compared with samples taken by the method being tested (see also Reference 4)
- performing systematic studies of sampling variables, such as the effect on the analytical results of mixing time, mixing method and time delay before analysis for samples of typical process material
- sampling prepared mixtures of known materials that simulate process material; this is frequently done in pilot plant studies to aid in the development of mixing and sampling systems and tentative operating procedures.

Tests of each sampling system should be repeated periodically to determine whether a change has occurred because material and equipment differences can affect mixing and sampling effectiveness. Further detail on techniques for monitoring sample quality is available in Reference 4 and the references cited in that report. The standard deviation of the estimate of sampling bias can be derived by statistical evaluations of the measurement results obtained with the test samples (see examples in Reference 40).

## 4.3 COMPARATIVE MEASUREMENTS

Parallel measurements by different measurement methods provides an alternative source of bias estimation. We refer to this approach as comparative measurements. Ideally, it would consist of remeasurement of samples or items by a method that is well established as an unbiased method for the application. In practical cases, the ideal method is often unavailable but acceptable alternatives can usually be found. Comparative measurement is a traditional method for substantiating chemical analyses for situations in which reference standards for calibration of the method are not available. Analytical methods based on different chemical and physical properties are not likely to have the same sources of bias. Therefore, agreement between results by two different methods increases confidence that the results from the method for routine use is unbiased. For some analyses, an alternative method known or widely accepted as free from bias is available for use as a standard method for calibration purposes although it may not be suitable for routine process application because of costs, slowness, large sample requirements, or other reasons.

Because suitable reference standards are available for the assay of SNM and their isotopes, it is possible to limit the use of comparative analyses for SNM to situations in which the measurements can be related to material standards or nationally accepted measurement systems. This is accomplished if the reference or alternative measurement is made by a method that is calibrated against suitable reference standards. To maintain traceability, a comparative measurement program may take any of these forms:(5)

- periodic submission of process samples for measurement by an alternative method whose results are traceable
- periodic submission of process samples for measurement by another facility having demonstrated traceability in the desired measurement
- interfacility interchange and measurement of well-characterized and representative materials whose values are assigned by a facility having demonstrated traceability in the measurement.

Comparative measurements can be used to monitor the accuracy of volume measurement instrumentation, using duplicate pneumatic gauges or manometers to check in-plant measurements of liquid level and using laboratory measurements to check in-plant solution density measurements. When comparative data indicate that a bias exists, an investigation should be initiated to verify the bias, identify the source and develop corrective measures.

NDA should also be monitored for bias by the comparative method.(14,17,18) NDA measurements are susceptible to bias resulting from differences in composition between unknown items and standards (see Section 3.5). It may not be feasible or practical to determine whether a given interfering condition exists in a given time, particularly for heterogeneous and variable material. In such cases, the comparative measurement method may be the only way to estimate the bias and its uncertainty. For example, even though a material is not well-characterized and it is not possible to predict all interferences and their behavior, conventional analyses (e.g., by a chemical assay method) of selected typical material can provide comparative data. The general approach would be to:

1. Periodically select an unknown and measure it several times with the NDA instrument to attain a mean value with a relatively small random error variance.
2. Measure the item using another technique for which calibration using reference standards is possible. This may be accomplished by:
    - dissolving the entire item and performing chemical analysis
    - dividing the item into a number of small samples and measuring each using a nondestructive small-sample assay system; the small sample NDA system should be one that has been calibrated using a calibrated destructive analysis method.

## 5.0 TECHNIQUES FOR MONITORING BIAS TEST DATA

To maintain control of the overall uncertainty of material balances within some limit, $L_0$, it is necessary to institute controls on the individual measurement processes such that their combined errors do not exceed $L_0$. The allowable errors of the individual measurement process may vary from situation to situation and ordinarily can be partitioned in several ways. To keep the overall uncertainty within $L_0$ requires that each particular measurement be controlled to within some limit, $L_i$ (i.e., that each measured value has a high probability of being within that limit). $L_i$ can be expected to differ considerably among measurement processes because of differences in the amount of SNM measured.

Because it is impractical to check each measurement, some procedure is selected for monitoring the measurement process performance that provides evidence that all but a small fraction of the measurements are within the control limit, $L_i$. Selection of the monitoring procedure should be based on a knowledge of the operating characteristics of the measurement process. These characteristics include random variances, systematic changes within the process, and the extent to which possible biases can offset all results from the true value. By sampling the measurement results at a sufficient rate and by an appropriate selection technique, reasonable assurance can be provided that the process is in control at a selected confidence level. The following paragraphs describe a few of the techniques which may be used as aids to monitoring errors.

Control charts provide a running graphical record of small subgroups of data taken from a repetitive process. The process may be a measurement process or a production process. Control charts may be kept for any of various statistical parameters of each small subgroup (e.g., the average, the standard deviation, or the range). The chart for each particular characteristic is designed to detect certain specified departures of the process from the assumed conditions. The order of groups is usually with respect to time, but not necessarily so. The grouping is such that the members of the same group are more likely to be alike than are members of different groups.(41)

Control charts can be used to test whether or not a process is in statistical control.(42,43) A process is in statistical control when repeated samples from the process behave as random samples from a stable probability distribution. Thus, the underlying conditions are such that it is possible to make a prediction in a probability sense.(41)

Control limits are usually computed by using formulas which utilize information obtained from the measurement results. The computed limits are placed as lines on a specific chart, and the process considered to be in control if all points fall within the control limits. These techniques are illustrated here with examples of $\bar{x}$ and Cusum charts.

An $\bar{x}$ control chart is illustrated in Figure 2 showing the percent deviation of individual analytical results from the accepted value of a reference standard. The control limits shown are based on the random error standard deviation of a single measurement and, in this case, two levels of control are indicated. For systems that are in a state of control, an occasional value slightly outside the inner limits, such as the three shown in Figure 2, is not unexpected (5 out of 100 can be expected, on the average, if the control limits are set at the 95% confidence level). Upon each such occurrence, a precautionary action may be to perform a repeat measurement. Two or more such points in

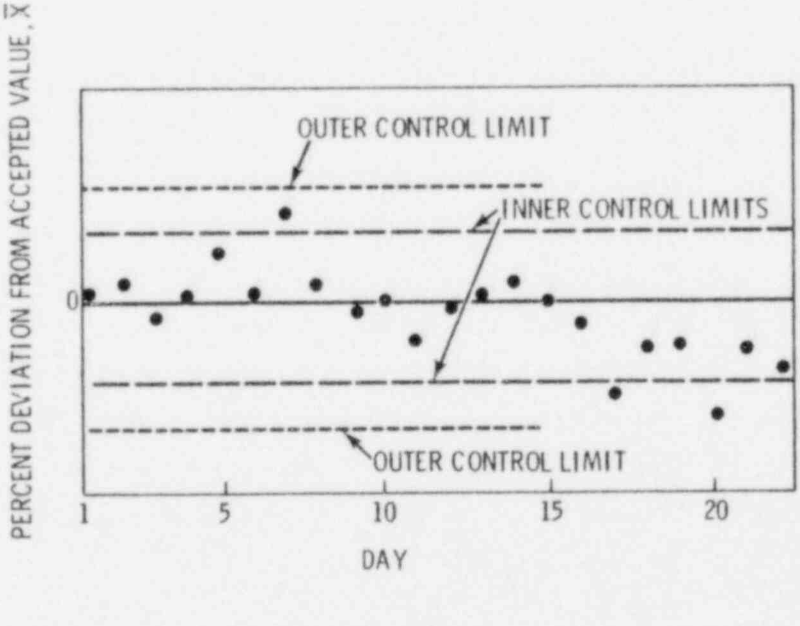

FIGURE 2. Results of Analysis of a Control Standard

close proximity or any result outside the outer limits should be cause for initiating a broader investigation. The number of data points above and below the mean should, on the average, be equal if the calibration of the system is stable and unbiased. Thus, a run of several consecutive values on one side of the mean, such as occurs in the week beginning on day 16 in Figure 2, indicates a possible shift in calibration. Such a hypothesis should be tested by an accepted statistical method.(40,41,44) More detailed discussions of the construction and use of control charts and other related statistical tests can be found in many References.(41-46)

A cumulative sum (Cusum) chart is a plot of consecutive sums of a series of differences of observed values or means of subsets of values from a reference or accepted value. For example, if $x_i$ is the result of the $i^{th}$ measurement on a reference standard and $\mu$ is the accepted value of this standard, the cumulative sum after the $k^{th}$ measurement is:

$$\text{Cusum} = \sum_{i=1}^{k} (x_i - \mu)$$

The series of values plotted on a Cusum chart would be:

$$\begin{array}{l} (x_1 - \mu) \\ (x_1 - \mu) + (x_2 - \mu) \\ (x_1 - \mu) + (x_2 - \mu) + (x_3 - \mu) \\ \qquad\qquad\qquad \vdots \quad \vdots \quad \vdots \quad \vdots \\ (x_1 - \mu) + (x_2 - \mu) \; . \; . \; . \; . \; + (x_{k-1} - \mu) + (x_k - \mu) \end{array}$$

If the summed values are means of subsets of measurements, then:

$$\text{Cusum} = \sum_{i=1}^{k} (\bar{x}_i - \mu)$$

One may choose to normalize the values to the standard deviation of x, in which case, the units on the ordinate of the chart would be in multiples of the standard deviation. The Cusum for this case is:

$$\sum_{i=1}^{k} (x_i - \mu)/\sigma_x$$

An illustration of a Cusum chart is given in Figure 3 in which the sum is normalized by dividing by the standard deviation of the x values. This chart also illustrates the use of a V-mask for outliers. The design and use of V-masks is described in References 47 and 48.

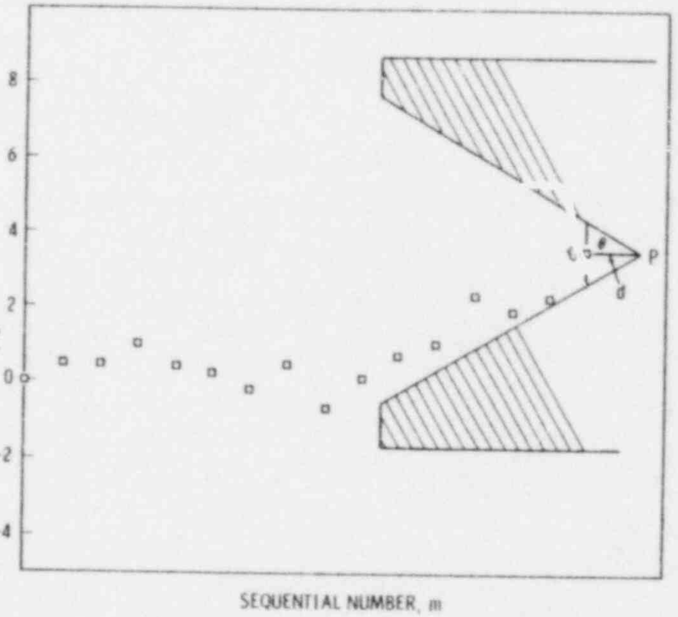

FIGURE 3. Typical Cumulative Sum Chart

Cusum charts provide a powerful graphical technique for detecting a persistant bias.(45,47,48) (If the bias is short-term or varies between positive and negative values, a standard Shewhart $\bar{x}$ chart may be just as effective for detecting the error.) If the measurement system is in statistical control and there is no bias relative to the reference value, $\mu$, the Cusum values are expected to be dispersed about a horizontal line with no persistent asymmetry or extreme departures from the average. An increased degree of dispersion relative to the immediately preceeding Cusum data points is indicative of increased random error or a recently introduced bias. A persistent positive or negative trend away from the expected horizontal (or zero) line is indicative of a persistent measurement bias. When there is no bias, all variables are random. Nevertheless, an occasional extreme deviate will shift the Cusum to a level appreciably above or below zero. The Cusum line will then appear to vary randomly about that level unless another large deviate of the opposite sign occurs to return the Cusum to near zero.

Several other graphical methods in addition to control charts, have been devised to detect biased data.(49,50) A few of these methods are described in the following paragraphs.

Results from the analysis of a group of standards may be plotted against the reference values in the manner shown in Figure 4. If there is no measurement bias, a straight line representing a good fit to the results should lie at 45° and pass through the origin. A constant bias would displace the line from the origin, and a bias that varies directly with the quantity of sample (proportionate bias) would cause the slope of the line to differ from 45°.

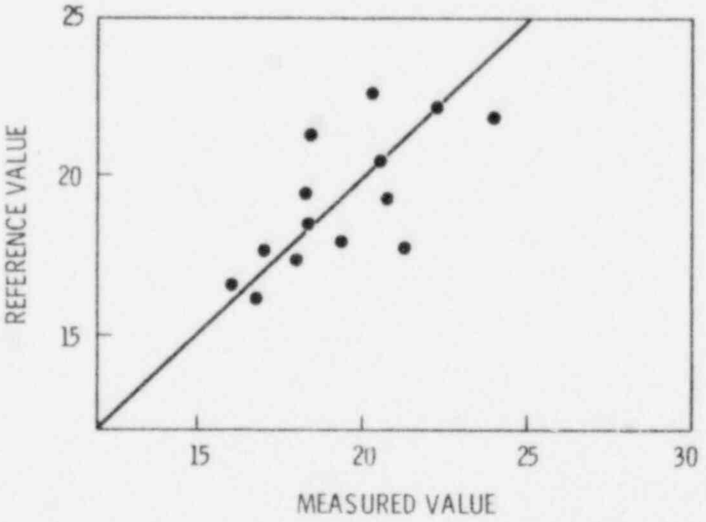

FIGURE 4. Analysis of a Standard

A similar graphical display is useful for comparing paired results by two measurement methods (see Figure 5). One of the methods is usually a standard procedure calibrated against reference standards and the test may be made with unknown or synthetic samples prepared to investigate specific effects such as chemical interferences. The data shown in Figure 5 indicate a bias between the methods. Such data can be statistically tested using a Student's t test. Similar graphs can be constructed for comparing results between instruments, analysts, or between results on different days.

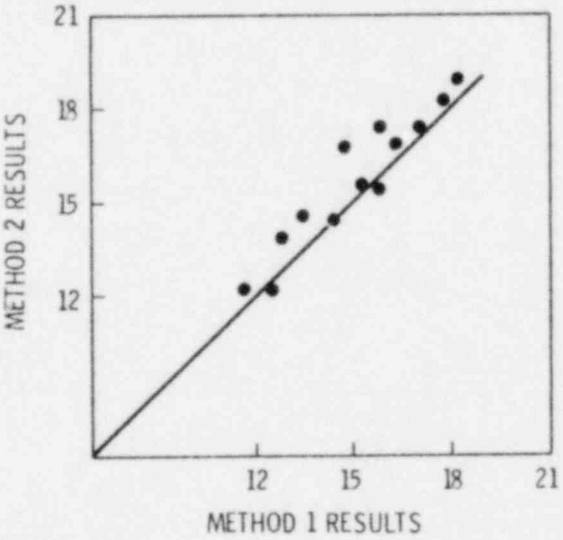

FIGURE 5. Analysis of Duplicate Samples by Two Procedures

A useful technique proposed by Youden(49,50) for comparing analytical results by several analysts or laboratories is displayed in Figure 6. Samples of two materials are analyzed for a given constituent by each participating analyst or laboratory and the results plotted as X and Y. A horizontal line is then drawn through the median or average Y value and a vertical line through the median or average X value. If there is no bias between analysts or laboratories, the results should be approximately equally distributed among the four quadrants formed by the median lines, reflecting simply random error. If a

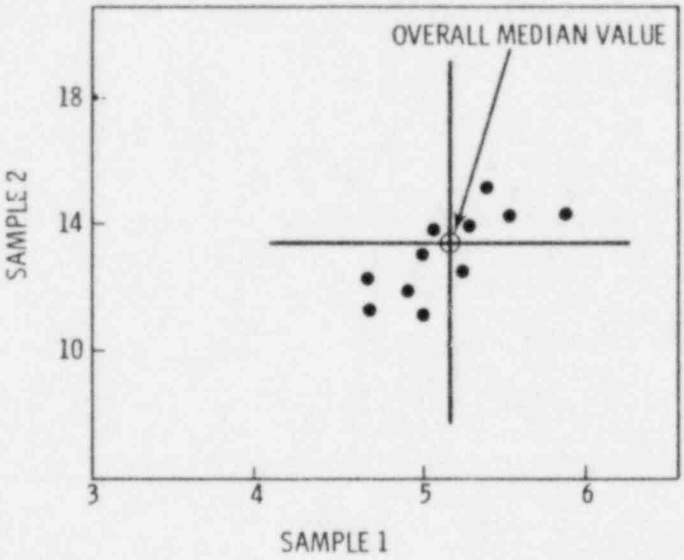

FIGURE 6. Analysis of Two Samples by Several Laboratories

bias between analysts or laboratories exists, the results will be found predominantly in the lower left and upper right quadrants as is the case in Figure 6.

Youden(50) has also suggested analyzing duplicate samples using different quantities of sample (e.g., using twice as much material for the second analysis as for the first) to test for a constant bias. When the two results are plotted against sample size, as in Figure 7, the line joining the results should pass through the origin (within the limits of error of the measurements). A constant bias in the measurements will cause an intercept at zero sample size. The intercept is an estimate of the bias. However, a proportionate bias will not be detected in this test, assuming that the true value of the constituent is not known.

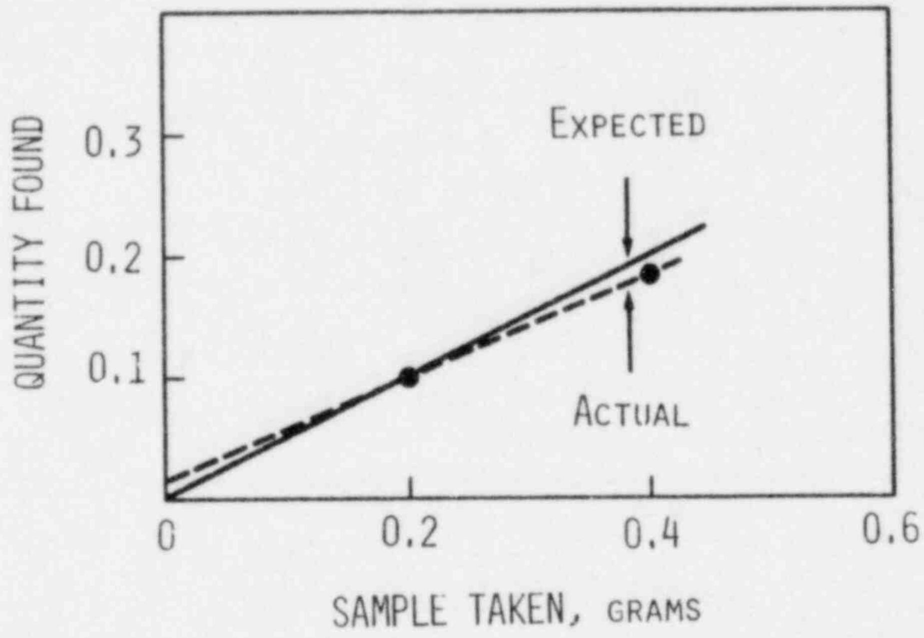

FIGURE 7. Comparison of Results at Two Levels

## 6.0 CONCLUSIONS

A knowledge of the statistical limitations of SNM control and accounting measurements is necessary to provide a quantitative basis for an alarm criterion that would initiate an investigation when discrepancies appear in material balances. To achieve this knowledge, all SNM measurement processes need to be characterized. This is done by obtaining information that describes 1) the measurement-to-measurement variations (random error behavior) of each process and 2) the relationship between the limiting mean of that process and some established and accepted frame of reference. This frame of reference could be a series of standards having values assigned by some authoritative organization, or it could be a highly characterized measurement process which has been accepted as a basis for comparison.

Once the random error standard deviation and bias associated with a measurement process have been estimated, the stability of the process must be demonstrated. This is necessary so that error estimates based upon measurement results obtained during a time interval such as a material balance period can be assigned with confidence to measurement results made at any time during the period. It is also necessary to monitor the measurement process to detect any changes in the magnitude of the associated errors.

Estimates of the bias of measurement processes are normally obtained from data generated by repetitive measurements of reference standards whose values are traceable to national standards or nationally accepted measurement systems. Supplementary data should be obtained (and in some cases, is imperative) by method testing and by comparative measurements with methods that can be independently calibrated. Systematic error behavior of all measurement processes should be continually monitored by one or more of these methods throughout the period of time that the method is applied to SNM accounting measurements. The limits of potential bias of measurement processes can be calculated from the calibration, method testing and comparative analysis data by the usual statistical methods.(25,29-31,40) It is also useful to monitor the results from measurements of reference standards with control charting techniques.

# REFERENCES

1. Code of Federal Regulations, "Title 10, Energy", Part 70, Section 70.57. Available from Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402, January 1977.

2. Ibid, Section 70.58.

3. R. J. Brouns, F. P. Roberts, J. A. Merrill and W. B. Brown. A Measurement Control Program for Nuclear Materials Accounting. NUREG/CR-0829, PNL-3021, Available from National Technical Information Service, Springfield, VA 22161, 1979.

4. R. J. Brouns, F. P. Roberts and U. L. Upson, Considerations for Sampling Nuclear Materials for SNM Accounting Measurements, NUREG/CR-0087, PNL-2592, Available from National Technical Information Service, Springfield, VA 22161, 1978.

5. Considerations for Establishing Traceability of Special Nuclear Material Accounting Measurements, Regulatory Guide 5.58, U.S. Nuclear Regulatory Commission Washington, DC 20555, 1978.

6. C. Eisenhart, "Realistic Evaluation of Precision and Accuracy of Instrument Calibration Systems," in Precision Measurement and Calibration, NBS SP-300, 1:21-47, H. H. Ku (ed), U.S. Government Printing Office, Washington, DC 20402, 1969.

7. J. T. Long, Engineering for Nuclear Fuel Reprocessing, Gordon and Breach, New York, NY, 1967.

8. Design Considerations - Systems for Measuring the Mass of Liquids, Regulatory Guide 5.48, U.S. Nuclear Regulatory Commission, Washington, DC 20555, 1975.

9. Conduct of Nuclear Material Physical Inventories, Regulatory Guide 5.13, U.S. Nuclear Regulatory Commission, Washington, DC 20555, 1973.

10. R. J. Brouns and F. P. Roberts, Procedures for Rounding Measurement Results in Nuclear Materials Control and Accounting. NUREG/CR-0033, PNL-2565, Available from Nuclear Regulatory Commission, Washington, DC 20555, 1978.

11. H. Harnby, "The Statistical Analysis of Particulate Mixtures, Part I, The Sampling of Mixtures and the Resultant Precision of Estimates," Power Technology, 5:81-86, 1972.

12. J. C. Williams and M. I. Kahn, "The Mixing and Segregation of Particulate Solids of Different Particle Size," The Chemical Engineer, 269:19-25, 1973.

13. Gutmacher et al., Methods for the Accountability of Plutonium Dioxide, USAEC Report WASH-1335, Available from National Technical Information Service, Springfield, VA 22161, 1974.

14. Nondestructive Assay of Special Nuclear Material Contained in Scrap and Waste, Regulatory Guide 5.11, U.S. Nuclear Regulatory Commission, Washington, DC 20555, 1973.

15. L. A. Kull, Catalogue of Nuclear Material Safeguards Instruments, USAEC Report, BNL-17165, Available from National Technical Information Service, Springfield, VA 22161, 1972.

16. Guide to Calibrating Nondestructive Assay Systems, ANSI N15.20-1975 Available from American National Standards Institute, 1430 Broadway, New York, NY 10018, 1975.

17. T. E. Shea, "Reduction, Control and Estimation of Nondestructive Assay Errors," in Proceedings of the 15th Annual Meeting, Institute of Nuclear Materials Management, Atlanta, GA, 1974.

18. J. E. Glancy, "Traceability of Nondestructive Assay to National Measurement System," in Proceedings of the 15th Annual Meeting, Institute of Nuclear Materials Management, Atlanta, GA, 1974.

19. Qualification, Calibration, and Error Estimation Methods for Nondestructive Assay, Regulatory Guide 5.53, U.S. Nuclear Regulatory Commission, Washington, DC 20555, 1975.

20. Mass Calibration Techniques for Nuclear Material Control, ANSI Standard N15.18, American National Standards Institute, 1430 Broadway, New York, NY 10018, 1975.

21. "Standard Methods for Testing Single Arm Balances," ASTM Standard E-319-68, Annual Book of ASTM Standards, Vol. 30, American Society for Testing and Materials, 1916 Race Street, Philadelphia, PA 19103.

22. T. W. Lashof and L. B. Macurdy, Precision Laboratory Standards of Mass and Laboratory Weights, NBS Circular 547, Section 1, Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402.

23. Volume Calibration Techniques for Nuclear Materials Control, ANSI Standard N15.19, American National Standards Institute, 1430 Broadway, New York, NY 10018, 1975.

24. Method for Liquid Calibration of Tanks, ASTM Standard D-1406-65, American Society for Testing and Materials, 1916 Race Street, Philadelphia, PA 19103, 1965.

25. S. C. Suda and D. F. Shepard, Effects of Measurement Errors on Linear Calibration Equations, BNL-50432, Available from National Technical Information Service, Springfield, VA 22151, 1976.

26. J. C. Hughes, Testing of Glass Volumetric Apparatus, NBS Circular 602, Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402, 1959.

27. Analytical Procedures for Accountability of Uranium Tetrafluoride, ANSI-N15.6-1972, and Analytical Procedures for the Accountability of Uranium Hexafluoride, ANSI-N15.7-1972, Available from American National Standards Institute, 1430 Broadway, New York, NY 10018, 1972.

28. W. R. Shields (ed), Analytical Mass Spectrometry Section; Instrumentation and Procedures for Isotopic Analysis, NBS Technical Note 277, National Bureau of Standards, Washington, DC 1966.

29. F. B. Stephens, R. G. Gutmacher, et. al., Methods for the Accountability of Uranium Dioxide, NUREG-75/010, Available from NTIS, Springfield, VA, 22161, 1975.

30. R. G. Gutmacher, F. Stephens, K. Ernst, J. E. Harrar, J. Magisted, T. E. Shea and S. P. Turel, Methods for the Accountability of Plutonium Nitrate Solutions, WASH-1282, Available from NTIS, Springfield, VA 22161, 1973.

31. F. B. Stephens, R. G. Gutmacher, K. Ernst and J. E. Harrar, Methods for the Accountability of Mixed Oxide, NUREG-0256, Available from NTIS, Springfield, VA 22161, 1977.

32. "Standard Recommended Practice for Calibration of Standards and Equipment," Standard D-2865-71, Annual Book of ASTM Standards, Volume 30, American Society for Testing and Materials, 1916 Race Street, Philadelphia, PA 19103, 1971.

33. Catalog of Standard Reference Materials, NBS Special Publication 260, Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402, 1973.

34. Calibration and Test Services of the National Bureau of Standards, NBS Special Publication 250, Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402, 1970.

35. E. Rein, S. F. Marsh, G. C. Swanson, G. R. Waterbury, G. L. Tietjen and R. K. Zeigler, Preparation of Working Calibration and Test Materials: Plutonium Nitrate Solution. NUREG-0118 (LA-NUREG-6348), Available from NTIS, Springfield, VA 22161, 1977.

36. S. S. Yamamura, F. W. Spraktes, J. M. Baldwin, R. L. Hand, R. P. Lash, and J. P. Clark, Preparation of Working Calibration and Test Materials: Uranyl Nitrate Solution, NUREG-0253, Available from NTIS, Springfield, VA 22161, 1977.

37. E. L. Garner, L. A. Machlin and W. R. Shields, Uranium Isotopic Standard Reference Materials, NBS Special Publication 266-27, U.S. Government Printing Office, Washington, DC 20402, 1971.

38. O. L. Davies, The Design and Analysis of Industrial Experiments, Second Edition, Hafner Publishing Co., 1967.

39. W. G. Cochran and G. M. Cox, Experimental Designs, John Wiley and Sons, 1950.

40. J. L. Jaech, Statistical Methods in Nuclear Material Control, TID-26298, Superintendent of Documents, U.S. Government Printing Office, Washington DC 20402, 1973, Chapter 3.

41. M. G. Natrella, Experimental Statistics, NBS Handbook 91, Superintendent of Documents, U.S. Government Printing Office, Washington, DC, 1963.

42. Control Chart Method of Controlling Quality During Production, American Standards Association Standard 21.3, American National Standards Institute 1430 Broadway, New York, NY 10018, 1958.

43. ASTM Manual on Quality Control of Materials, Part 3, ASTM Special Technical Publication 15-C, American Society for Testing and Materials, 1916 Race Street, Philadelphia, PA 19103, 1951.

44. W. G. Cochran, Sampling Techniques, Chapter 1, 2nd Edition, John Wiley and Sons, Inc, New York, NY, 1963.

45. A. J. Duncan, Quality Control in Industrial Statistics, 4th Edition, R. D. Irwin, Inc., Homewood, Ill., 1974.

46. A. Hald, Statistical Theory with Engineering Applications, John Wiley & Sons, 1952, pp. 338-373.

47. N. L. Johnson and F. C. Leone, Cumulative Sum Control Charts, Part I, Industrial Quality Control, July 1962.

48. E. S. Page, "Cumulative Sum Charts," Technometrics, 3, No. 1, pp. 1-9, 1961.

49. W. J. Youden, "Graphical Diagnosis of Interlaboratory Test Results," Page 133 in Precision Measurement and Calibration, NBS Special Publication 300, Vol. 1, Harry H. Ku (ed), Superintendent of Documents, U.S. Government Printing Office, Washington DC, 1969. (See also Technometrics, 1, 409 1959.)

50. W. J. Youden, "The Sample, The Procedure and The Laboratory," NBS Special Publication 300, Page 138. (See also Analytical Chemistry, 32, No. 13, 1960.)

NRC FORM 335
(7-77)

U.S. NUCLEAR REGULATORY COMMISSION

**BIBLIOGRAPHIC DATA SHEET**

1. REPORT NUMBER *(Assigned by DDC)*
NUREG/CR-1284
PNL-3264

4. TITLE AND SUBTITLE *(Add Volume No., if appropriate)*
Methods of Determining and Controlling Bias in Nuclear Material Accounting Measurements

2. *(Leave blank)*

3. RECIPIENT'S ACCESSION NO.

7. AUTHOR(S)
R. J. Brouns, J. A. Merrill, F. P. Roberts

5. DATE REPORT COMPLETED
MONTH January
YEAR 1980

9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS *(Include Zip Code)*
Pacific Northwest Laboratory
Battelle Memorial Institute
P.O. Box 999
Richland, WA 99352

DATE REPORT ISSUED
MONTH June
YEAR 1980

6. *(Leave blank)*

8. *(Leave blank)*

12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS *(Include Zip Code)*
Office of Standards Development
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555

10. PROJECT/TASK/WORK UNIT NO.

11. CONTRACT NO.
B20149

13. TYPE OF REPORT
Technical Report

PERIOD COVERED *(Inclusive dates)*

15. SUPPLEMENTARY NOTES

14. *(Leave blank)*

16. ABSTRACT *(200 words or less)*

The common sources of bias associated with measurements made for nuclear control and accounting purposes are described and recommended approaches to monitoring and controlling such errors are presented. Measurement bias is placed in perspective through a disussion of the effect of measurement error in material balance and accounting. A recommended program for monitoring and controlling bias includes calibrations, routine measurements of reference standards, method testing and comparative analyses.

17. KEY WORDS AND DOCUMENT ANALYSIS

Bias
Sources of bias
Measurement error
Measurement calibrations

17a. DESCRIPTORS

17b. IDENTIFIERS/OPEN-ENDED TERMS

18. AVAILABILITY STATEMENT
Unlimited

19. SECURITY CLASS *(This report)*
Unclassified

20. SECURITY CLASS *(This page)*
Unclassified

21. NO. OF PAGES

22. PRICE
$