

NUREG/CR-0830
PNL-3022

Monitoring the Random Errors of Nuclear Material Measurements

Prepared by R.J. Brouns, F.P. Roberts, J.A. Merrill

Pacific Northwest Laboratory

Operated by
Battelle Memorial Institute

**Prepared for
U.S. Nuclear Regulatory
Commission**



NOTICE

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights.

Available from

GPO Sales Program
Division of Technical Information and Document Control
U. S. Nuclear Regulatory Commission
Washington, D. C. 20555

Printed copy price: \$2.25

and

National Technical Information Service
Springfield, Virginia 22161

Monitoring the Random Errors of Nuclear Material Measurements

Manuscript Completed: June 1980
Date Published: June 1980

Prepared by R.J. Brouns, F.P. Roberts, J.A. Merrill

Pacific Northwest Laboratory
P.O. Box 999
Richland, WA 99352

**Prepared for Division of Siting, Health and Safeguards Standards
Office of Standards Development
U.S. Nuclear Regulatory Commission
Washington, D.C. 20555
NRC FIN No. B20149**



ABSTRACT

Monitoring and controlling random errors is an important function of a measurement control program. This report describes the principal sources of random error in the common nuclear material measurement processes and the most important elements of a program for monitoring, evaluating and controlling the random error standard deviations of these processes.

SUMMARY

A measurement control program monitors the accuracy and precision of measurement processes and provides controls for selection and application of measurement methods to help ensure consistent and acceptable measurement quality.⁽¹⁾ Licensees subject to the requirements of Section 70.58 of 10 CFR⁽²⁾ must establish and maintain a measurement control program for special nuclear material (SNM) accounting measurements. Close control of SNM measurement quality is needed to ensure that a loss, theft, or diversion of SNM will not be masked by either bias or excessive random error in the measurement data used in the periodic material balances that are performed to verify the adequacy of SNM control and accounting. This report describes the principal sources of random error and the technical and administrative elements considered to be important in monitoring and controlling random error standard deviations of SNM accounting measurements.

Estimates of the random error standard deviations of the measurement processes are usually derived from data obtained by performing repeated independent measurements of typical process materials and items and by analyzing duplicate samples. Because it is important for the standard deviation estimates to be representative of typical routine SNM accountability measurements, the repeated and/or duplicate measurements should be:

- made in the same manner as routine measurements, i.e., the same equipment, instruments, operators, and procedures
- made in an independent manner to ensure that all sources of random error are just as probable in duplicate measurements as they are in isolated measurements
- distributed evenly over the period of time that the SNM accountability measurements are made.

The data collection functions of measurement control of random error must be supported by an adequate statistical evaluation effort to provide experimental designs, measurement control data analyses, hypothesis tests,

and determination of limits for control charts. In addition, the organization of the nuclear facility should be structured to facilitate effective administrative control over:

- review and formal approval of designs, specifications and installations of measurement and sampling equipment and other equipment
- review and approval over the selection of measurement methods, reference standards and calibration procedures
- control of selection, training and qualifying personnel who perform SNM measurements and related tasks, such as sampling and calibrations.

A periodic audit should be performed to ensure continued adherence by all personnel to the established measurement control program policies and procedures.

CONTENTS

ABSTRACT	iii
SUMMARY	v
1.0 INTRODUCTION	1
2.0 SOURCES OF RANDOM ERROR	3
2.1 VOLUME MEASUREMENTS	3
2.2 MASS MEASUREMENTS	4
2.3 SAMPLING	5
2.4 CHEMICAL AND ISOTOPIC ANALYSES	5
2.4.1 Chemical Analyses	5
2.4.2 Isotopic Analyses	7
2.5 NONDESTRUCTIVE ASSAY (NDA)	9
3.0 MONITORING AND CONTROLLING RANDOM ERROR	13
4.0 EVALUATION AND MONITORING OF RANDOM ERROR VARIANCE COMPONENTS	17
4.1 ESTIMATING RANDOM ERROR	17
4.2 COMBINING ERRORS AND CALCULATING LEID	18
4.3 PLANNING MEASUREMENT CONTROL DATA COLLECTION	18
4.4 NUMBER OF DEGREES OF FREEDOM REQUIRED	18
4.5 STATISTICAL DESIGN OF EXPERIMENTS AND ANALYSIS OF VARIANCE	19
4.6 REJECTION OF OUTLIER DATA	20
4.7 EVALUATING DATA TRENDS	21
REFERENCES	23
APPENDIX A	A-1

1.0 INTRODUCTION

This report describes the procedures that are recommended for monitoring the random error variances of measurement systems for SNM accounting. The report was prepared by the Pacific Northwest Laboratory (PNL)^(a) for the Safeguards Standards Branch of the Office of Standards Development, Nuclear Regulatory Commission.

The concept of a measurement system implies that the measurement process can be well characterized and can be maintained in a state of control. A measurement system can be said to be operating in a state of control when:

- 1) sequences of independent replicate measurements support a single valued limiting mean, 2) the collection of measurement results is free from obvious trends or groupings, and 3) each new measurement result verifies the validity of prediction limits based on historical performance data.⁽³⁾

The inherent variability of measurement processes is the principal source of the inventory difference in a material balance when there have been no net losses or gains of material. A knowledge of the characteristics of the measurement processes, primarily their accuracy and precision, is the necessary first step in controlling the uncertainty of each measurement process within acceptable limits and in maintaining a statistically-based alarm criterion for initiating corrective action if the SNM inventory difference becomes excessive.

To continually characterize and control measurement processes, a measurement control program consisting of a system of frequent replicate measurements of process materials, standard measurements, calibrations, comparative analyses, and data evaluation techniques should be maintained.⁽¹⁾ The functions of the program are to continually monitor the state of statistical control of all key measurement processes and to control and, if possible, improve the quality of the SNM accounting measurements.

A measurement control program should be maintained in support of nuclear materials measurements as a matter of prudent management. In addition, the Code of Federal Regulations, Title 10, Part 70, requires licensees authorized

(a)PNL is operated by Battelle Memorial Institute.

for possession and use of more than one effective kilogram of special nuclear material (SNM) to maintain a measurement control program (see Sections 70.57(a) and 70.58(f) of Reference 2). Included in the program requirements is the statement, "The program shall generate current data on the performance of measurement processes, including, as appropriate, values for bias corrections and their uncertainties, random error variances, limits for systematic errors, and other parameters needed to establish the uncertainty of measurements pertaining to materials control and accounting." (See Section 70.57(b)(8) of Reference 2.) This report is concerned specifically with random error.

Random error may be described as the chance variation encountered in all measurement work, characterized by the random occurrence of both positive and negative deviations from a mean value, the algebraic average of which approaches zero with a sufficiently large number of measurements. A random error behaves as if it were chosen at random from a population of such errors having a given probability density function. Its associated density function expresses how often the different values of the random variable will occur.⁽⁴⁻⁶⁾ By contrast, a systematic error is a constant, unidirectional component of error that affects all members of a data set; its value can usually be estimated by the deviation of the limiting mean of a measurement process from a reference value. A systematic error is commonly referred to as a bias and it can often be corrected for. The magnitude of the random error of a mean value is reduced by replicate measurements, whereas this is not true for systematic errors.

2.0 SOURCES OF RANDOM ERROR

2.1 VOLUME MEASUREMENTS

Volume measurements are usually made in calibrated tanks equipped with instruments for determining the liquid level and the liquid level reading is converted to a volume estimate with a calibration curve or equation. The calibration equation of a linear volume measurement tank is $h_L = \alpha + \beta V$, and the working equation for volume measurement is $V = (h_L - \alpha)/\beta$, where V is the volume of liquid, h_L is the liquid height in the tank, and α and β are calibration constants, namely, the intercept and slope of the calibration curve, respectively. 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 the density of the liquid and the pressure difference between a point near the bottom of a tank and a point in the tank above the surface of the liquid. The liquid density is needed to convert pressure differential measurements to estimates of liquid level. It also may be measured by a pneumatic probe system.

The mathematical function for measurements of differential pressure with a liquid-filled manometer, which is a very common method in nuclear processing facilities is:

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

where h_M is the liquid height in the manometer and ρ_M and ρ_L are the densities of the liquids in the manometer and the tank, respectively. From this it is clear that random error can occur in reading the manometer, h_M , and in measuring the tank liquid density, ρ_L . An error in the density value for the manometer liquid, ρ_m , would be a bias since it would be constant for repeated tank level measurements as long as a fixed value of ρ_M is used. ρ_M is, in essence, a calibration constant.

The random errors of volume measurements are due primarily to the inherent variability of instrumental measurements and to the variability introduced by operators reading liquid height and density instruments. Random errors may also be introduced by variations in ambient temperature that change the density of the manometer fluid and by incomplete mixing of tank contents, which can affect the density measurements. In addition, cloudy manometers and poor lighting can contribute to operator error in reading manometers. Other liquid level measurement systems, such as sight glasses, floats and sonic instruments are subject to similar sources of random error.

2.2 MASS MEASUREMENTS

The usual sources of random error in weighings are the variations associated with the inherent limits of reproducibility of the scale itself and the variations contributed by the operators.^(a) When weight differences are used, such as gross weights minus tare weights, the random errors of the gross weighing and the tare weighing both contribute to the random error of the net weight. The random error of weighing also can be affected by data rounding procedures.^(6,7)

Although the reproducibility of a weighing device is fixed primarily by its design and construction, the actual precision achieved in use is affected by the mechanical condition of the scale and environmental circumstances such as vibration and temperature fluctuations. The mechanical condition can deteriorate through wear, corrosion, accumulated dirt at critical points, and mechanical damage to critical parts. In addition, the manner in which a weight is placed on a platform scale, such as off-center loading, may cause a change in lever ratio that has an effect on the reading scale.⁽⁸⁾

The random error contributed by operators can be affected by the design of the scale or balance readout system. Digital readout and direct printing of the weight are advantageous because they reduce operator-contributed reading variations and reduce the likelihood of mistakes.

(a)The term scale will be used to refer to any mass measurement device.

2.3 SAMPLING

The usual source of random error in sampling is the variability of material composition within the lot or batch sampled. Sampling errors can also be caused by changes in sample composition during sampling or while awaiting analysis, and by contamination of sample containers or sampling equipment.

Large sampling errors may occur with heterogeneous materials even if multiple sample increments are taken and composited to reduce the error. Powder mixtures that tend to segregate during mixing operations according to particle size or density,⁽⁹⁾ liquid-solid mixtures, scrap, and contaminated waste materials are particularly hard to sample representatively.

It should be recognized that an appreciable sample variability can also occur in splitting or subsampling primary or gross samples to obtain suitable increments for analysis if the gross sample is not first thoroughly blended. Coarse materials should be preground to a sufficiently small particle size for subsampling.⁽¹⁰⁾ The magnitude of sampling error may be affected by deficiencies in the procedures or by the mixing and sampling equipment, as well as by operator carelessness.

The composition of samples can change appreciably during the splitting, subdividing or grinding, blending and subsampling stages if materials that are susceptible to evaporation, or moisture or oxygen gain or loss, are not adequately protected from adverse environments. Such changes may usually lead to systematic error but can also contribute to random error (see Reference 10 for additional detail and references).

2.4 CHEMICAL AND ISOTOPIC ANALYSES

2.4.1 Chemical Analyses

The chemical analysis methods in common use for SNM accountability are:

- Isotopic Dilution Method for U and Pu. Used for uranium and plutonium determinations when the concentrations are very low or the quantity of samples is restricted, such as with highly radioactive material.

- Titration Methods for U and Pu. Uranium and plutonium are usually determined by chemical or coulometric titration methods when the sample size is not limited. The preferred amount of element in samples for these methods ranges from about 10 mg to a few grams. These are the most precise methods available with the exception of the gravimetric method for uranium.
- Gravimetric Method for U. Applied to pure uranium compounds and solutions. (Gravimetric methods that require a purification step, such as precipitation followed by gravimetry, are seldom used in the nuclear industry.)
- Alpha Counting for Pu. Applied to materials with very low concentrations of plutonium.
- Fluorometric Method for U. Applied to materials with very low concentrations of uranium.
- Spectrophotometric Method for U. Applied to materials with low and intermediate concentrations of uranium.

Some recommended sources of additional detail about the analytical methods for uranium and plutonium are References 11 through 15.

There are additional analytical methods that affect SNM accounting to some degree but their sources of error are sufficiently similar to the generic sources discussed here that further detail is unwarranted. Examples are the emission and x-ray spectrometry methods used for determinations of minor constituents that may interfere with uranium and plutonium determinations and the chemical pretreatment methods for instrumental measurements such as x-ray and mass spectrometry.

Chemical analyses of materials involve many measurement steps and manipulations, each having the potential for introducing error into the overall measurement. Typical steps are drying, mixing, subsampling, weighing, dissolving, chemically separating or purifying and measuring composition-related parameters by chemical, physical or instrumental techniques. Common sources of random error in these operations are:

1. Nonhomogeneity of samples, which may cause subsampling error.
2. Variability in the weight and volume measurements of analytical samples and reagents.
3. Variability of titration endpoints (usually related to variability in the endpoint reaction and the indicator or endpoint sensor response).
4. Variability in reaction stoichiometry or rate (may be due to variations in temperature, reaction time, amount of reagents added and impurities).
5. Variability in the response of measurement instruments, such as may be caused by meter or chart recorder dead zone, electronic instability, and transient signals or "noise" from the detector (e.g., amperometric electrodes) or power supply.
6. Variability in the amount of interfering constituents in samples.
7. Variations induced by the analyst, such as quantities or concentrations of reagents and reaction times.

2.4.2 Isotopic Analyses

The preferred method for the isotopic analysis of heavy elements is surface ionization mass spectrometry, although for some applications, other methods are also used. For example, gamma spectrometry is applicable to a limited number of isotopes and is used where a rapid isotopic determination or a nondestructive technique is needed and a poorer precision can be accepted. Alpha spectrometry is usually used for determining ^{238}Pu in plutonium at isotopic concentrations below about 0.1%. In addition, gas source mass spectrometry is usually used for the ^{235}U enrichment determinations of UF_6 feed material for fuel fabrication and for the uranium isotopic ratio measurements in enrichment facilities. The gas source method is capable of better precision than the surface ionization method but has a more limited applicability in isotopic analyses of nuclear materials.

The discussion of errors in isotopic analysis will focus on the surface ionization method because of its much wider application in nuclear materials

accounting measurements. An abundance of literature is available on isotopic analysis. Some suggested sources for the surface ionization method are References 11, 12, 16 and 17.

The isotopic analysis procedure for surface ionization mass spectrometry consists of: 1) chemical separation and purification of the desired element, e.g., uranium or plutonium; 2) mounting, drying and calcining a measured portion of a dilute solution of the purified element directly on a ribbon filament or pair of filaments to be inserted in the mass spectrometer ion source (rhenium metal filaments are used for uranium and plutonium samples); and 3) analyzing the metal ion spectra emitted from the source at high filament temperatures ("surface" ionization or "thermal" ionization) by energy/mass resolution in the mass spectrometer. A variety of mass spectrometers are used, differing primarily in ion source and ion collector designs. Single and triple filament sources, Faraday trap and electron multiplier collectors and d.c. and pulse counting readouts are the common variations. Mass spectrometers are among the most complex analytical instruments used and considerable operational knowledge and skill are required to achieve the accuracy and precision needed for SNM material balance accounting.

Random errors in isotopic analysis affect the standard deviation of the isotopic ratio. Random variability is observed when measuring ratios repeatedly on a single filament and when analyzing two or more filaments prepared from the same sample. Standard deviations calculated from multiple filament data are more representative of the actual uncertainty of an isotopic analysis because they account for both within-filament and between-filament effects. The sources of variability that may contribute to the random error of an isotopic analysis are:

- Instrument variability, which is a function of the stabilities of the electronic components, source and detector noise, and the effect of power supply influences such as voltage transients. (Instrument stability is usually high because a high degree of stability and self-regulation is built into the mass spectrometer electronics.)

- Isotopic fractionation, which results in a change in the isotopic ratios over the period of the analysis. The rate is affected by the filament temperature, sample composition, quantity of sample, and the mounting and pretreatment procedures. (Fractionation may cause both bias and random error, depending on the way the operator takes data.)
- The number of repeated measurements on a single filament.
- The signal intensity, if the intensity is low enough that counting statistics or peak height measurement uncertainty is an appreciable source of random error.
- Chart recorder dead zone and linearity, if d.c. measurement of the detector signal is recorded graphically.
- Short-term variations in instrument bias.
- Differences in sample filaments, such as quantity, purity and composition of the element and the loading and treatment procedure.
- Differences between operators in the manner in which they process the samples and collect the data.
- Differences in source and collector noise levels, which may be affected by pressure, cleanliness, memory and electronic effects.

The chemical procedures that precede the isotopic analysis may affect the results because, as stated previously, sample purity, acidity and concentration of the element can affect filament characteristics such as evaporation rates, ionization efficiencies, and source discrimination. Errors stemming from poor sample quality control may be avoidable by an experienced mass spectrometer operator because the sample mounting or ionization characteristics observed when the analysis is begun will be noticeably abnormal and the analysis can be rejected and another sample obtained.

2.5 NONDESTRUCTIVE ASSAY (NDA)

Nondestructive assay of special nuclear materials is based on measuring either fission-induced or spontaneously occurring radiation from the isotopes of uranium, plutonium, their decay products, or some combination of these. ⁽¹⁸⁻²⁰⁾

Nondestructive assay techniques permit rapid analysis without affecting the material assayed. They also permit assay of heterogeneous materials without sampling and chemical processing, thereby avoiding potentially large sampling errors.

Error in NDA measurements is known to be affected by the: 1) operational characteristics of the measurement system (both geometrical and electronic), 2) uniformity and energy of the stimulating radiation (for active NDA) and the signal and background radiation, 3) the amount and distribution of SNM and other materials in the sample containers, and 4) the uniformity in composition and dimensions of the containers.⁽¹⁸⁻²²⁾ The errors in NDA can occur both from measurement to measurement of the same item and from item to item within a single category of material.^(19,21) A distinction between these types of error is useful because the former can be routinely determined by replicate assays of items but the latter may require independent or verifying measurements of the items by another assay technique.

Measurement-to-measurement error is caused primarily by random variables in the measurement process, although systematic error can also be a cause. For example, an instrumental instability may be manifested in sensitivity drift in one direction, resulting in biased results, or in short-term changes that result in random error. Item-to-item error may be random over a group of items but the statistical distribution of the error cannot be assumed to be normal and the mean may not be zero because the error may be caused by factors that result in bias for any finite set of items. For example, an error due to a particular SNM distribution in the unknown items may be all in one direction because the mean characteristic distribution of SNM in the calibration standards used is appreciably different from the mean distribution in the unknowns.

Item-to-item error will not be treated further in this report because the normal methods of random error standard deviation measurement do not evaluate such errors. They are regarded, for our purposes, as short-term systematic errors in the sense discussed by Jaech⁽²³⁾ and their evaluation is achieved by calibrations, comparative measurements and other bias determination procedures.

The principal sources of random, measurement-to-measurement error in NDA are:

- inherent variability in the number of counts observed (counting statistics)
- instrument or electronic variability
- ambient environmental (temperature and humidity) and voltage fluctuations
- fluctuations in the background radiation
- variations in item positioning with respect to detectors and, if applicable, with respect to the excitation source (this source of random error may be operator dependent)
- changes in distribution of SNM or matrix material within a container, such as the settling or shifting of high-density material.

3.0 MONITORING AND CONTROLLING RANDOM ERROR

Random error standard deviations are usually determined from data obtained by replicate sampling and measurement of process materials. In the case of sampling and chemical assay, a standard deviation determined in this manner represents the combined sampling and analytical sources of error. In order to determine the magnitude of the individual standard deviations, it is necessary to perform duplicate measurements on the replicate samples. The random error standard deviation is calculated from the differences between the replicate measurement results. Sets of duplicate measurements are often used and the statistical analysis of the pair differences gives the standard deviation.⁽²⁴⁾

To ensure that an error estimate represents the behavior of routine measurements, all routine procedural steps and conditions should be followed in the replicate measurements and the measurements should be made independent of each other to the extent possible. Thereby, every potential source of random error has the same chance of affecting each result. For example, when duplicate samples are taken to determine sampling error, independent selection of sample duplicates may entail choosing at random the samples or items, the time of sampling, and the person taking each sample. Normally, if two samples are taken in sequence from a lot, they are likely to be taken by the same individual in an identical manner under identical circumstances (e.g., temperature, humidity, etc.) and from portions of the lot that were processed at about the same time. Therefore, the difference between duplicate results would tend to be smaller than it would be if the samples were taken independently. Similar precautions need to be taken to achieve independent replication of bulk measurements and analyses. Some laboratories attempt to have remeasurements run "blind" (disguised) to avoid the possibility that the random error estimate may be affected by knowledge of the original results or by knowledge that it is a test sample. Another common practice is to have the duplicate measurement made on another work shift or on another day. This may not affect complete independence of measurements unless disguising is also accomplished but it will incorporate between-operator and between-day effects into the variance estimate.

Data from research and development studies on a measurement method and from the analysis of standards may be used to estimate the precision of a method and to identify its sources of error. Measurements of precision obtained in research and development studies are useful for predicting method performance but may not give true estimates of the random error variance affecting routine applications because development data are usually generated by a few select persons, on a single set of apparatus, in a research laboratory, and/or with standards and synthetic samples.

Standard deviation also may be estimated from replicate measurements of reference standards, but it is common for the estimates to be smaller than those calculated from replication of routine process measurements. The probable reason is that in the case of standards, precision is improved by the uniformity of composition and the natural tendency of analysts to exercise greater care when analyzing a standard than when analyzing routine samples. Consequently, standard deviation estimates from standards data should not be assumed to be representative of measurement performance in routine process measurements.

Since the random error behavior of sampling and measurements is not necessarily constant over time, it should not be assumed that data collected in the past are applicable to current measurements.^(a) Also, to ensure that error data reflects current conditions, monitoring should be frequent and should be distributed fairly evenly over the time during which the measurements are applied. In addition, the dynamic character of a measurement system requires that the error data collection plan be flexible. Conditions affecting the measurement processes such as material composition, equipment used for sampling or measurements, personnel, and procedures can vary from time to time, necessitating revision of the error data collection plan. A continuous evaluation of data requirements can assure that the current needs are being met and that unneeded data are not being collected. No universal pattern for acquiring measurement error data can be prescribed that would be effective for all types of facilities. Each has unique characteristics that can affect

(a) This places constraints on data pooling. Despite the advantage of pooling, which is to reduce the variance of the estimated standard deviation by increasing the number of degrees of freedom, sets of error monitoring data should not be pooled without confirmation by statistical tests that the variances of the sets are the same.

the kind and/or amount of data needed to evaluate measurement uncertainties. The sources of error in measurements, the magnitude of error in terms of quantity of SNM, and the potential impacts on the limit of error of inventory differences (LEID) need to be considered, also, in establishing the amount of effort devoted to error monitoring. The statistical aspects of planning data collection and the number of degrees of freedom that should be acquired for monitoring random error are discussed in Sections 4.3 and 4.4.

The data collection and evaluation functions of a measurement control program should be supported by an organization that provides some additional control functions to facilitate high measurement quality. In this respect the program should provide for:

- reviews and formal approval controls over design, specifications and installation of measurement and sampling equipment
- review and approval controls over the selection of measurement methods, reference standards, and calibration procedures
- control of the selection, training, and qualifying of personnel who perform SNM measurements and related tasks such as sampling and calibrations
- periodic audits by management to determine whether the approved program policies and procedures are implemented and effectively carried out.

The following are examples of specific practices that should be maintained to facilitate effective control over the magnitude and variability of measurement random errors:

1. When scales and other measurement instruments are read manually, operators should be trained to read, interpolate, and round off data in a consistent and uniform way. A scale designed to minimize parallax and interpolation error should be used. Digital reading and printing scales should be used, if feasible.
2. The measurement control program should provide for periodic inspection, testing, and maintenance of scales and other measurement instruments by a qualified instrument specialist. Formal procedures and inspection

frequencies should be developed for each instrument. The frequency for each instrument should be consistent with its performance history and other circumstances, including the severity of service conditions.

3. Sampling procedures should be tested prior to acceptance as a standard practice so that procedural and equipment deficiencies can be corrected in a timely manner. The tests should establish optimum mixing methods, optimum mixing times, and, for certain sampling equipment, optimum flushing times. The tests should be repeated when control data indicate a need to reduce sampling error or when process or equipment changes are made that may alter sample representiveness.
4. When it is necessary to sample materials that do not lend themselves to blending, such as scrap and heterogeneous solid wastes, special sampling procedures should be developed that avoid unacceptably large sampling errors. Techniques such as pulverization and incineration of the material before sampling should be used, if practical. Scrap, wastes and similar material should be batched or packaged by material type, source, composition, and isotopic content to minimize variability among items for NDA or to reduce sampling error for chemical analysis.
5. Since the measurement error for a sample includes laboratory subsampling and sample preparation errors, in addition to analytical error, samples should be thoroughly blended before subsampling and sample preparation for analysis, and the adequacy of the blending procedure should be known from tests. The procedure for determining the standard deviation of analytical measurement results should be designed to include subsampling and sample preparation effects. If separation of the errors due to the various sources is desired, a statistically designed experiment similar to that described in Appendix A is necessary.
6. The results of random error data evaluation should be summarized periodically by the measurement control group and the random error monitoring data, test results, procedure and equipment changes, and maintenance should be documented. Finally, control charts should be used to provide a visual record of random error monitoring data for control of the performance of each measurement method.

4.0 EVALUATION AND MONITORING OF RANDOM ERROR VARIANCE COMPONENTS

Analyses and interpretations of measurement data involve estimating variances of individual SNM measurements and accumulations of measurements such as material balances, limits for rejecting outliers, and control limits for monitoring the precision and accuracy of measurements. These topics are discussed in many texts on the statistical analysis of data. (6, 25-27)

4.1 ESTIMATING RANDOM ERROR

The random error variance of a measurement is estimated from replicate measurement data using the formula:

$$s^2 = \frac{\sum_{i=1}^n (x_i - \bar{x})^2}{n-1}$$

for n measurements, x_i , of one sample or item, or using

$$s^2 = \frac{\sum d_i^2 - (\sum d_i)^2/n}{2(n-1)}$$

for n sets of duplicate values, where $d_i = x_{1i} - x_{2i}$ and s^2 is the variance estimate. For the case where the expected value of the difference is zero, i.e., there is no bias between the duplicates, the second equation reduces to:

$$s^2 = \sum d_i^2 / 2n.$$

The variance estimate should be obtained from measurements made under conditions like those prevailing during routine use of the measurement process to avoid omission of sources of error. For example, if it is the normal practice to use several instruments or several operators for the process measurements, the estimates calculated from replications are applicable to all measurements only

if the replicate data involve all of these sources of error, unless tests have shown that between-instrument and between-operator effects are negligible.

4.2 COMBINING ERRORS AND CALCULATING LEID

The standard deviations of shipments, receipts and ID's, which are sums or linear combinations of many individual measured values, are of primary interest in SNM accounting. Calculation of the standard deviations of these quantities requires consideration of the standard deviations of calibration, sampling, analysis, bulk measurement, and bias correction. A system of statistical procedures for propagating such errors and calculating LEID (twice the standard deviation of the inventory difference) is explained and illustrated with practical examples by Jaech.⁽⁶⁾

4.3 PLANNING MEASUREMENT CONTROL DATA COLLECTION

Establishing a data collection schedule for monitoring samples and measurements requires a knowledge of: 1) the measurement methods being used and their expected precisions, 2) the processing schedules and material flows, and 3) any planned changes in either processes or measurement methods. When individual measurement errors are combined to obtain LEID, the statistician performing the calculations usually knows how much data are available for each error source estimate, how current the data are, and what trends in the various sources of error have occurred. In addition, he can assess the relative contribution of the various sources of error in the overall material balance. From this information the statistician can identify measurement deficiencies and plan future replicate samplings and analyses, and other tests of measurement quality in a manner that ensures adequate control data but avoids the acquisition of excessive information at unnecessary expense.

4.4 NUMBER OF DEGREES OF FREEDOM REQUIRED

Confidence limits for an estimated standard deviation depend on the number of degrees of freedom (d.f.) (usually one less than the number of measurements in the case of simple replicate measurements) associated with the calculated standard deviation. Tests of statistical significance of an

apparent bias or an apparent change of standard deviation are related to the confidence interval. (Discussions of confidence interval determinations and hypothesis testing are found in most statistics books.)^(6,25,26) The confidence interval for an estimate of variance, s^2 , as a function of the number of degrees of freedom in the estimate can be estimated from a table of the χ^2 function where:

$$\chi^2 = \frac{s^2}{\sigma^2} (\text{d.f.}).$$

Such an analysis shows that appreciable gains in the limits for s^2 are achieved by increasing the degrees of freedom up to about 15 as illustrated in Figure 1. As a general guide, therefore, the minimum number of sets of duplicate measurements for determining the random error standard deviation of a measurement process should be about 15.

The number of replicate sets of measurements in each situation and their spacing over time to adequately monitor random error during a given time interval is largely a matter of statistical judgment. The judgment concerns the number of repeated measurements or analyses needed to ensure that there is an adequate number of degrees of freedom to obtain a sufficiently precise estimate of the appropriate variance component.

4.5 STATISTICAL DESIGN OF EXPERIMENTS AND ANALYSIS OF VARIANCE

Special studies or experiments often are necessary to develop improved measurement processes and to identify and estimate sources of error when random error has multiple causes. One example is to estimate the standard deviations due to sampling and due to analysis separately. Such studies are greatly facilitated if they are appropriately designed. The design and analysis of a three-factor nested experiment, from which can be estimated separately the standard deviations between containers, samples, and analyses, is given in Appendix A. Examples of experimental designs useful in measurement control are described in References 27 and 28.

CONFIDENCE LIMITS ON s

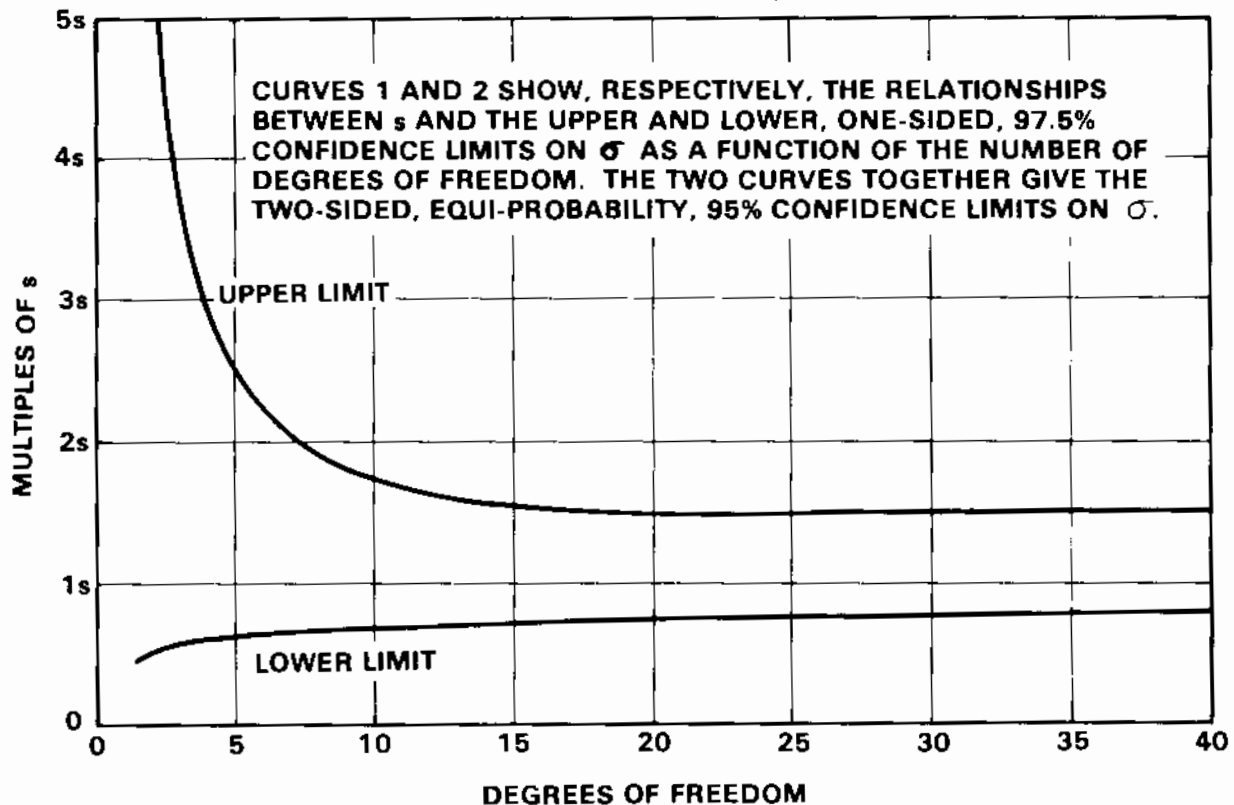


FIGURE 1. 95 Percent Confidence Limits on σ as a Function of Degrees of Freedom

4.6 REJECTION OF OUTLIER DATA

A common statistical problem in measurements is dealing with outlier observations. An outlier may be merely an extreme manifestation of the ordinary variability in the data, in which case the value should be retained. On the other hand the outlier may have been caused by an identifiable physical deviation in a procedure, by a miscalculation, or by an error in data recording or transcribing. The first step in dealing with an outlier is to look for a physical cause; if one is established the outlier should be rejected. If no physical cause is found or if there is any doubt about the cause, a statistical test should be applied.

For a single outlying observation one test is to determine whether the difference between the outlier and the mean of all of the data (including

the outlying value) is within acceptable limits about the mean, based on an estimate of the variance of the current data. Acceptable procedures for setting the limits for rejecting an outlier have been published by ASTM,⁽²⁹⁾ F. E. Grubbs,⁽³⁰⁾ G. L. Tietjen, and R. H. Moore.⁽³¹⁾

4.7 EVALUATING DATA TRENDS

The sequences of results from the analysis of ranges of replicate measurements or estimates of standard deviation should be monitored to detect significant changes in magnitude, particularly increases. Control charts are convenient devices for observation of such sequential data. When a shift or trend is suspected, a physical explanation for the apparent shift is sought. If there is no physical explanation, statistical tests should be applied to the data. Hypothesis testing with χ^2 tests or F tests is a method for detecting a significant change in the variance of a measurement process (see References 6, 25, 26). Prior to a hypothesis test, it is often desirable to test for normality and non-randomness of the underlying data. (See Chapter 2 in Reference 6.)

Control charting procedures help monitor measurement quality by providing operators, analysts and measurement control personnel with a practical and rational means of detecting loss of statistical control.⁽³²⁾ The use of control charts is similar to hypothesis testing. A control chart is usually a sequential plot of data with pre-established lines indicating the expected value and the upper and lower control limits. The control limits are based on previous estimates of variability for the measurement process while it was in statistical control; the 2σ and 3σ limits are the most common.

The usual control charts for monitoring standard deviations or other measures of dispersion are range or R charts and standard deviation or s charts. An R or s chart should be used with each measurement process. Figure 2 is an example of a range chart for paired data (R_2). A useful reference containing the details for constructing R charts and s charts is Quality Control and Industrial Statistics by A. J. Duncan.⁽³⁴⁾

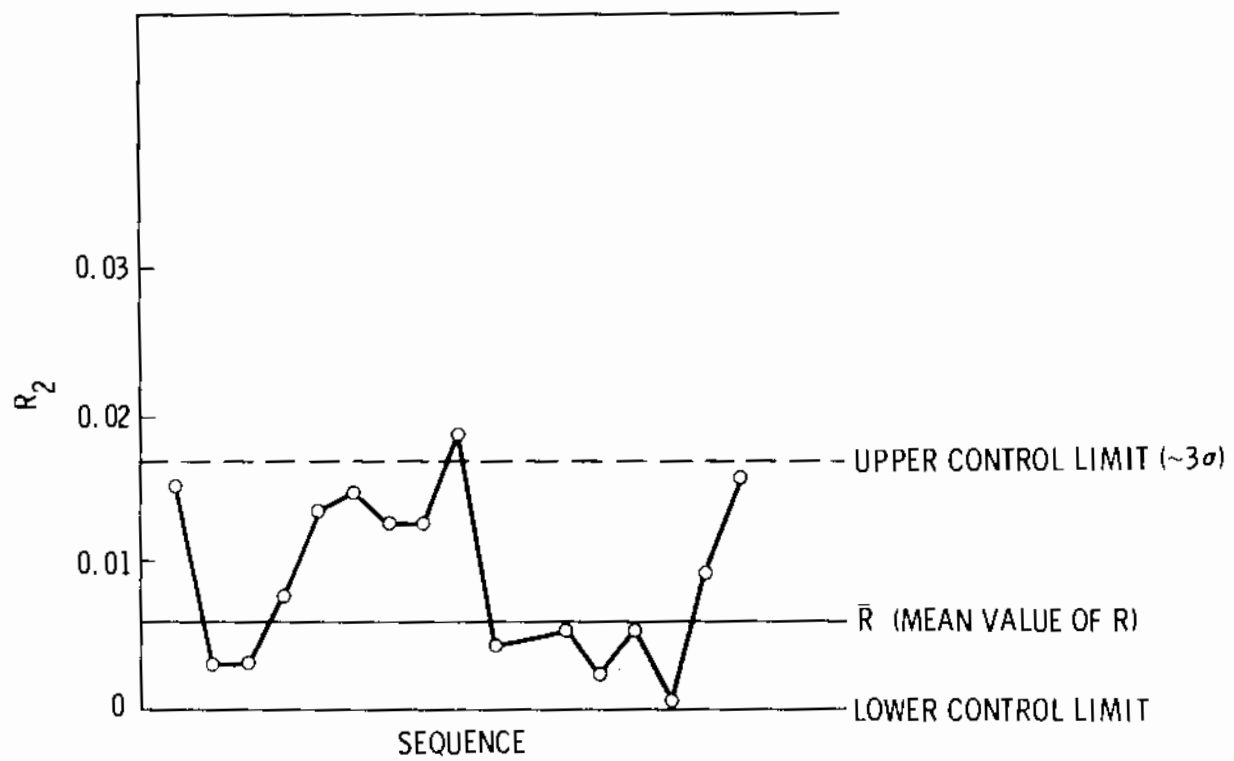


FIGURE 2. A Chart for the Range of Two Values

REFERENCES

1. 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), National Technical Information Service, Springfield, VA 22161, 1979.*
2. Code of Federal Regulations, Title 10, Energy. Superintendent of Documents U.S. Printing Office, Washington DC, January 1979.
3. P. E. Pontius, Notes on the Fundamentals of Measurement and Measurement as a Production Process. National Bureau of Standards Report NBS IR 75-545, 1974.
4. Safeguards Dictionary. WASH-1173, U.S. Atomic Energy Commission, July 1, 1971 (available from Superintendent of Documents, U.S. Government Printing Office, Washington DC).
5. Statistical Terminology and Notation for Nuclear Materials Management. ANSI N15.5-1980, American National Standards Institute, 1430 Broadway, New York, NY 1980.
6. J. L. Jaech, Statistical Methods in Nuclear Material Control. TID-26298, Superintendent of Documents, U.S. Government Printing Office, Washington DC, 1973.
7. R. J. Brouns and F. P. Roberts, Procedure for Rounding Measurement Results in Nuclear Materials Control and Accounting. NUREG/CR-0033, (PNL-2565), National Technical Information Service, Springfield, VA 1978.*
8. Mass Calibration Techniques for Nuclear Material Control. ANSI N15.18-1975, American National Standards Institute, 1430 Broadway, New York, NY, 1975.
9. J. C. Williams and M. I. Khan, "The Mixing and Segregation of Particulate Solids of Different Particle Size." The Chemical Engineer, 269:19-25, January 1973.
10. R. J. Brouns, F. P. Roberts and U. L. Upson, Considerations for Sampling Nuclear Materials for SNM Accounting Measurements. NUREG/CR-0087 (PNL-2592), National Technical Information Service, Springfield, VA, 1978.*
11. C. J. Rodden (ed), Selected Measurement Methods for Plutonium and Uranium in the Nuclear Fuel Cycle. TID-7029 (2nd Ed.), National Technical Information Service, Springfield, VA, 1972.
12. F. B. Stephens, R. G. Gutmacher, K. Ernst and J. E. Harrar, Methods for the Accountability of Mixed Oxide. NUREG-0256, National Technical Information Service, Springfield, VA, 1977.*

13. R. G. Gutmacher, F. Stephens, K. Ernst, S. P. Turel and T. E. Shea, Methods for the Accountability of Plutonium Dioxide. WASH-1335, Superintendent of Documents, U.S. Government Printing Office, Washington DC, 1974.
14. R. G. Gutmacher, F. Stephens, K. Ernst, J. E. Harrar, J. Magistad, T. E. Shea and S. P. Turel, Methods for the Accountability of Plutonium Nitrate Solutions. WASH-1282, Superintendent of Documents, U.S. Government Printing Office, Washington DC, 1974.
15. F. B. Stephens, R. G. Gutmacher, K. Ernst, J. E. Harrar and S. P. Turel, Methods for the Accountability of Uranium Dioxide. NUREG-75/010 National Technical Information Service, Springfield, VA, 1975.*
16. E. L. Garner, L. A. Machlan and W. R. Shields, Uranium Isotopic Standard Reference Materials. NBS SP-26D-27, Superintendent of Documents, U.S. Government Printing Office, Washington DC, 1971.
17. W. R. Shields (ed), Analytical Mass Spectrometry Section: Instrumentation and Procedures for Isotopic Analysis. NBS Technical Note 277, Superintendent of Documents, U.S. Government Printing Office, Washington DC, 1966.
18. L. A. Kull, Catalogue of Nuclear Material Safeguards Instruments. BNL-17165, Brookhaven National Laboratory, Upton, NY, 1972. (NTIS).
19. Nondestructive Assay of Special Nuclear Material Contained in Scrap and Waste. Regulatory Guide 5.11, U.S. Nuclear Regulatory Commission, Washington DC, October 1973.
20. T. D. Reilly and M. L. Evans, Measurement Reliability for Nuclear Material Assay. LA-6574, National Technical Information Service, Springfield, VA, 1977.*
21. T. E. Shea, "Reduction, Control and Estimation of Nondestructive Assay Errors," Proceedings of the 15th Annual Meeting of the Institute for Nuclear Materials Management. Atlanta, GA, June 19-21, 1974 (Nucl. Material Management, Vol. III, No. 3, Fall 1974 pp. 341-350), INMM, Kansas State University, Manhattan, KS, 1974.
22. P. Ting, Verification of Prior Measurements by Nondestructive Assay. NUREG-0230, National Technical Information Service, Springfield, VA, 1977.*
23. J. L. Jaech, Statistical Methods in Nuclear Material Control. Chapter 3.
24. J. L. Jaech, Statistical Methods in Nuclear Material Control. Chapter 8.

25. C. A. Bennett and N. A. Franklin, Statistical Analysis in Chemistry and the Chemical Industry, Chapter 8, John Wiley and Sons, Inc., New York, NY 1954.
26. J. Mandel, The Statistical Analysis of Experimental Data, Interscience Publishers, John Wiley and Sons, Inc., New York, NY, 1967.
27. F. H. Tingey, "Design of Experiments in Analytical Chemistry Investigation." in Treatise on Analytical Chemistry, Part 1, Vol. 10, I. M. Kolthoff, and P. J. Elving (eds.), John Wiley and Sons, Inc., New York, NY, 1972.
28. N. J. Youden, "Critical Evaluation of an Analytical Procedure." in Encyclopedia of Industrial Chemical Analysis. Vol. 1, p. 755, Snell and Helton, Interscience Publishers, New York, NY, 1966.
29. Standard Recommended Practice for Dealing with Outlying Observations ASTM Standard E-178-68. American Society for Testing and Materials, 1916 Race Street, Philadelphia, PA, 1968.
30. F. E. Grubbs, "Procedures for Detecting Outlying Observations in Samples." Technometrics, 11(1):1-12, 1969.
31. G. L. Tietjen and R. H. Moore, "Some Grubbs-Type Statistics for the Detection of Several Outliers." Technometrics, 14:583, 1972.
32. Control Chart Method of Controlling Quality During Production. ANSI Z1.3-1958, American National Standards Institute, 1430 Broadway, New York, NY, 1958.
33. A. J. Duncan, Quality Control and Industrial Statistics, (4th Ed.), Chapter 21, Richard D. Irwin, Inc., Homewood, IL, 1974.

* Also available for purchase from the NRC/GPO Sales Program,
U.S. Nuclear Regulatory Commission, Washington, D.C. 20555

APPENDIX A
EXPERIMENTAL DESIGN FOR SEPARATING THREE RANDOM ERROR VARIANCES: A
THREE-FACTOR NESTED DESIGN

It is sometimes necessary to determine the random error standard deviations attributable to individual sources of error such as different work shifts, different instruments, and different analysts or operators. For example, the variance calculated from analyses of replicate samples represents the combined errors of sampling and analysis. When it is necessary to isolate sampling error, appropriate patterns of replicate analyses, e.g., replicate analyses of sets of replicate samples, are usually designed to separate analytical and sampling variances, using appropriate statistical methodology.^{(24-29)(a)} Jaech has given several examples of experiments designed to determine the effect of individual sources of error of measurements in the nuclear processing field.⁽²⁴⁾ Generally, such experiments involve replicate measurements of unknowns, standards, or both under selected sets of conditions, so arranged that an "analysis of variance" can be performed.

An experimental design based on an analysis of variance is commonly used to determine variances due to various treatments, measurement conditions, etc. In this example, assume one wished to determine the standard deviations due to random analytical error, random sampling error, and random variation between containers. Assume that the determination can be made by selecting two samples at random from each of ten containers of material and that each sample is analyzed in duplicate. The outcome would be 40 analytical results representing 2 samples from each of the 10 containers.^(b) The data can be arranged as shown in Table A-1.

The data collected with this pattern of measurements permit calculation of the overall mean value for the 10 containers based on 40 analyses, the mean for each container based on four analyses; and the mean for each sample based on two analyses. The calculations are generally summarized in an analysis of variance table such as Table A-2.

(a)The references are listed at the end of the report.

(b)This would often be an insufficient amount of data, but this case is simplified for illustrative purposes.

TABLE A-1. Schematic of Measurements Based
on a Balanced Nested Design

<u>Containers</u>	<u>Samples</u>	<u>Analyses</u>
1	1	X_{111}, X_{112}
	2	X_{121}, X_{122}
2	1	X_{211}, X_{212}
	2	X_{221}, X_{222}
.		
p	1	X_{p11}, X_{p12}
	2	X_{p21}, X_{p22}

TABLE A-2. Analysis of Variance^(a)

Source of Estimate	Degrees of Freedom	Sums of Squares ^(b)	Mean Squares	Average Value of Mean Square ^(c)
Between Containers	p-1	$SS_i = nq \sum_1^p (x_i - \bar{x})^2$	$MS_c = SS_i / (p-1)$	$\sigma_a^2 + n\sigma_s^2 + nq\sigma_c^2$
Between Samples Within Containers	p(q-1)	$SS_{j(i)} = n \sum_{j=1}^{pq} (x_{ij} - \bar{x}_i)^2$	$MS_s = SS_{j(i)} / p(q-1)$	$\sigma_a^2 + n\sigma_s^2$
Between Analyses Within Containers and Samples	pq(n-1)	$SS_{k(ij)} = \sum_{k=1}^{pqn} (x_{ijk} - \bar{x}_{ij})^2$	$MS_a = SS_{k(ij)} / pq(n-1)$	σ_a^2
TOTAL	pqn-1	$SS = \sum_{k=1}^{pqn} (x_{ijk} - \bar{x})^2$		--

where

σ_a^2 = analytical variance

σ_s^2 = sampling variance

σ_c^2 = container variance

x_{ijk} = the kth analysis of the jth sample of the ith container,

or $x_{ijk} = \mu + c_i + s_{j(i)} + a_{k(ij)}$

with $i = 1, 2, \dots, p$
 $j = 1, 2, \dots, q$
 $k = 1, 2, \dots, n$

and μ = the overall process mean

c_i = the container effect

$s_{j(i)}$ = the sample effect within container

$a_{k(ij)}$ = the analysis effect within sample and container

(a) Assumes a large population and a balanced nested experimental design.

(b) Simpler formulas for the calculations can be derived.

(c) Note that the factors n, p, and q may need to be adjusted in the case of finite population of containers, samples per container, or analyses per sample.

The three variances are estimated from these data by comparing sums of squared differences between the observed results and the mean of each group. It can be shown that:

1. The mean sum of squared difference between duplicate analyses estimates the analytical variance, σ_a^2 .
2. The mean sum of squared differences between sample means and container means estimates the combination of the analytical and sampling variances, $\sigma_a^2 + n\sigma_s^2$, where n is the number of analyses per sample (two in this case).
3. The mean sum of squared differences between container means and the overall mean of 10 containers estimates the combination of the analytical, sampling, and container variances, $\sigma_a^2 + n\sigma_s^2 + nq\sigma_c^2$, where q is the number of samples per container (two in this case).

The mean squares are equated with the proper variance functions, and the equations are simply solved algebraically to provide estimates of the analytical, sampling, and container variances. The estimates of the three variance components are:

$$\begin{aligned}\hat{\sigma}_a^2 &= MS_a \\ \hat{\sigma}_s^2 &= \frac{MS_s - MS_a}{n} \\ \hat{\sigma}_c^2 &= \frac{MS_c - MS_s}{nq}\end{aligned}$$

In practice, it is possible for either (or both) of $\hat{\sigma}_s^2$ or $\hat{\sigma}_c^2$ to turn out negative. In such cases, an approach that is frequently used is the following:

- If $MS_s < MS_a$,

then set $\hat{\sigma}_s^2 = 0$

$$\hat{\sigma}_a^2 = \frac{SS_{j(i)} + SS_{k(i,j)}}{p(nq-1)}$$

$$\text{and } \hat{\sigma}_c^2 = \frac{MS_c - \hat{\sigma}_a^2}{nq} \text{ if greater than or equal to zero.}$$

If this estimate for $\hat{\sigma}_c^2$ is less than zero, then set

$$\hat{\sigma}_s^2 = \hat{\sigma}_c^2 = 0$$

$$\text{and } \hat{\sigma}_a^2 = \frac{SS}{pqn-1}$$

- If $MS_s > MS_a$ and $MS_c < MS_s$,

$$\text{then set } \hat{\sigma}_a^2 = MS_a$$

$$\hat{\sigma}_c^2 = 0$$

$$\text{and } \hat{\sigma}_s^2 = \frac{\frac{SS_i + SS_{j(i)}}{pq-1} - MS_a}{n} \text{ if greater than or equal to zero.}$$

If this estimate for $\hat{\sigma}_s^2$ is less than zero, then set

$$\hat{\sigma}_s^2 = \hat{\sigma}_c^2 = 0$$

$$\text{and } \hat{\sigma}_a^2 = \frac{SS}{pqn-1}$$

Also, statistical tests can be applied using the information in the analysis of variance table to test the hypothesis of $\hat{\sigma}_s^2 = 0$ and the hypothesis of $\hat{\sigma}_c^2 = 0$. The hypothesis of $\hat{\sigma}_s^2 = 0$ is tested using the F test given by

$$F = \frac{MS_s}{MS_a}, \text{ with } p(q-1) \text{ and } pq(n-1) \text{ degrees of freedom.}$$

The hypothesis of $\hat{\sigma}_c^2 = 0$ is tested, using

$$F = \frac{MS_c}{MS_s}, \text{ with } p-1 \text{ and } p(q-1) \text{ degrees of freedom.}$$

If the calculated ratio is greater than the tabulated critical F- value at the chosen significance level (the 0.05 or the 0.01 level is typically used), then the corresponding hypothesis is rejected.

The above experimental pattern and analysis of variance is quite simple. However, the principles of the process may be used to determine the variances of methods of analysis, different analysts, and different environmental conditions, in addition to the factors used in the example.

NRC FORM 335 (7-77)		U.S. NUCLEAR REGULATORY COMMISSION BIBLIOGRAPHIC DATA SHEET		1. REPORT NUMBER (Assigned by DDC) NUREG/CR-830	
4. TITLE AND SUBTITLE (Add Volume No., if appropriate) Monitoring The Random Errors of Nuclear Material Measurements				2. (Leave blank)	
7. AUTHOR(S) R.J. Brouns, F.P. Roberts, and J.A. Merrill				3. RECIPIENT'S ACCESSION NO. 5. DATE REPORT COMPLETED MONTH: June YEAR: 1980	
9. PERFORMING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code) Pacific Northwest Laboratory P.O. Box 999 Richland, WA 99352				DATE REPORT ISSUED MONTH: YEAR:	
12. SPONSORING ORGANIZATION NAME AND MAILING ADDRESS (Include Zip Code) Office of Standards Development U.S. Nuclear Regulatory Commission Washington, DC 20555				6. (Leave blank) 8. (Leave blank)	
10. PROJECT/TASK/WORK UNIT NO. MP 474-7				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) Monitoring and controlling random errors is an important function of a measurement control program. This report describes the principal sources of random error in the common nuclear material measurement processes and the most important elements of a program for monitoring, evaluating, and controlling the random error standard deviations of these processes.					
17. KEY WORDS AND DOCUMENT ANALYSIS Statistics Systematic Error Statistical Methods Standard Deviations Random Error Measurement Control Program			17a. DESCRIPTORS Volume Measurements Mass Measurements Estimating Random Error Mass Measurements		
17b. IDENTIFIERS/OPEN-ENDED TERMS Chemical Analyses, Isotopic Analyses, Nondestructive Assay (NDA), Inventory Difference (ID), Limit of Error of ID (LEID)					
18. AVAILABILITY STATEMENT			19. SECURITY CLASS (This report)		21. NO. OF PAGES
			20. SECURITY CLASS (This page)		22. PRICE \$

