NBS SPECIAL PUBLICATION 676-I

U.S. DEPARTMENT OF COMMERCE/National Bureau of Standards

Measurement Assurance Programs Part I: General Introduction

Measurement Assurance Programs
Part I: General Introduction

Brian Belanger

Office of Physical Measurement Services
National Measurement Laboratory
National Bureau of Standards
Washington, DC 20234

U.S. DEPARTMENT OF COMMERCE, Malcolm Baldrige, Secretary
NATIONAL BUREAU OF STANDARDS, Ernest Ambler, Director

Issued May 1984

Library of Congress Catalog Card Number: 84-601030

National Bureau of Standards Special Publication 676-I
Natl. Bur. Stand. (U.S.), Spec. Publ. 676-I, 71 pages (May 1984)
CODEN: XNBSAV

U.S. GOVERNMENT PRINTING OFFICE
WASHINGTON: 1984

For sale by the Superintendent of Documents, U.S. Government Printing Office, Washington, DC 20402

Foreword

This two-part guide has been written in response to numerous inquiries concerning the use of Measurement Assurance Program (MAP) Services offered in conjunction with the Calibration Services of the National Bureau of Standards. MAP Services differ from the usual Calibration Services because they focus on the quality of measurements being made by the participant rather than just the properties of the participant's instruments or standards. The services are offered as an adjunct to the user's own measurement control programs and are designed to help a laboratory quantify the uncertainty of its measurements relative to national standards. MAP services represent a new approach to "calibration" and this publication presents the rationale for their use.

This guide consists of two parts published separately:

- Part I - A General Introduction, authored by Brian C. Belanger, Chief, Office of Physical Measurement Services.
- Part II - Development and Implementation, authored by Carroll Croarkin, Statistical Engineering Division.

Part I of this guide is intended as a general introduction to MAPs. It is not intended as a specification for the kind and magnitude of the effort required to ensure that measurements are adequate for their intended use. It is intended to provide enough detail on NBS MAP Services to help potential users decide whether or not such services can play a useful role in their measurement activities. The first part of this guide gives illustrations of proven approaches to measurement control to assist the reader in constructing or upgrading an internal quality control program.

Part II is concerned with the development and implementation of MAPs. Particular emphasis is placed on principles for statistical analysis and interpretation of MAP data, including characterization of measurement errors, use of control charts and specific examples of MAPs in process.

Users of these MAP Services must determine what constitutes adequate measurement quality control for their applications. Not every laboratory will find it necessary or desirable to use NBS Measurement Assurance Program Services. This guide will enable users of NBS MAP Services to ensure that the services are utilized more effectively.

George A. Uriano
Director
Measurement Services
National Bureau of Standards
April 15, 1984

TABLE OF CONTENTS

LIST OF FIGURES

Measurement Assurance Programs

Part I. General Introduction

1. Introduction and Purpose

This is the first of a two-part guide. This part explains the concept of measurement quality assurance, describes the services offered by the National Bureau of Standards (NBS) to support measurement quality assurance, and provides information concerning the benefits that could accrue to users of NBS Measurement Assurance Program (MAP) services and those employing measurement quality control procedures. Part II describes the statistical techniques used to implement measurement assurance programs.

It is useful to distinguish between a Measurement Assurance Program (MAP) and NBS MAP services. A MAP is the measurement quality assurance program implemented by the participant outside NBS to ensure accurate measurements relative to national standards. NBS MAP services are offered by the Bureau to aid in the achievement of measurement quality control in the participating laboratory and to link the measurements in that laboratory to national standards.

This publication (the first part of the guide) does not attempt to cover the detailed technical material needed to establish and maintain a MAP in any of the technical areas mentioned. A bibliography is provided for the reader interested in understanding the MAP approach as applied to any specific technical area. A glossary of terms related to measurement assurance is provided to help the reader understand this and other publications on measurement assurance.

The MAP concept is still evolving. MAP services that NBS currently offers have been developed somewhat independently of each other; consequently, differences exist among these services. As the approach to measurement quality control evolves, both the concept and its application will probably change. NBS MAP services may become more uniform in their approach and design, and new MAP services will certainly incorporate the best features of existing services. Thus, it is important that NBS receive feedback from users of its MAP services.

The MAP approach to measurement quality control (QC) or quality assurance (QA) for physical measurements is not particularly unique*. Those familiar with the principles of QC and applied statistics will recognize that most, if not all, of the features of a MAP are tools that are well known in the QA field. In fact, to a great extent, a MAP can be thought of as statistical quality control procedures developed many years ago by Shewhart and others applied to a measurement process**. Similar methods have been used to ensure the accuracy of industrial chemical processes, clinical laboratory and biological laboratory measurements, etc. The essential feature of a MAP is that it focuses on the whole physical measurement process: the operator, the environment, the methods, in

---

* The terms "quality control" and "quality assurance" are defined in the glossary.

** The term "measurement process" is defined in the glossary.

addition to the instrument. The purpose of a MAP is to establish relative to national standards the uncertainty of the measurements being made, and to monitor that uncertainty on a continuing basis to ensure that the measurements are sufficiently accurate for their intended application.

For those who want only an "executive summary" of the subject, the next section consists of answers to the 19 questions most frequently asked about MAPs.

2. The Nineteen Most Frequently Asked Questions About MAPs: An Executive Summary

1. WHAT IS A MAP?

A MAP is a quality assurance program for a measurement process that quantifies the total uncertainty of the measurements (both random and systematic components of error) with respect to national or other designated standards and demonstrates that the total uncertainty is sufficiently small to meet the user's requirements*.

2. HOW DOES AN NBS MAP SERVICE DIFFER FROM AN NBS CALIBRATION SERVICE?

NBS MAP services focus on the quality of measurements being made in the participating laboratory rather than on the properties of the participants instruments or standards. Conceptually, participation in a MAP service can be thought of as a way of "calibrating" the entire laboratory.

In an NBS calibration, the customer's device or standard is sent to NBS to be calibrated. When the device or standard is returned to the customer, the customer receives an NBS test report containing measured value(s) of the standard and an associated measurement uncertainty(s) relative to national standards. The uncertainty reported on an NBS calibration is a measure of the quality of the NBS calibration process and is not a property of the instrument or standard or the customer's measurement system.

The proper use of a calibrated standard can result in accurate measurements in the customer's laboratory. However, if the operators are not sufficiently skilled, if the environmental conditions of the laboratory differ from those at NBS, if unsound measurement procedures are used, or if other problems (known or unknown) exist, then the measurements actually made in the customer's laboratory may not be nearly as accurate as the uncertainty of the NBS-calibrated standard would, in principle, permit. Without some comparison between the NBS measurement process and the customer's, no unequivocal statement can be made about the actual accuracy of the laboratory's measurements.

The MAP service quantifies the total uncertainty of the participant's measurement process. In order to establish this uncertainty, it is necessary for the participating laboratory to

---
* The terms uncertainty, random error, and systematic error are defined in the glossary.

have an ongoing measurement control program. In such a program measurements are repeated on one or more stable standards in order to estimate the random error associated with the partcipating laboratory's measurement process.

In a typical MAP service, a stable artifact (or set of artifacts) referred to as a "transport standard,"* is measured at NBS and sent to a participating laboratory for measurement by that laboratory. The value of the transport standard is normally unknown to the participant. Following measurements by the participant, the transport standard is returned to NBS for remeasurement. The NBS data and the data from the participating laboratory are then analyzed, and a test report is sent from NBS to the participant stating the offset of the participating laboratory's measurement process from national standards and the total uncertainty of the participant's calibration process.

The total uncertainty of the participating laboratory's measurement process reflects both the random error (a measure of the reproducibility, precision, or within-laboratory variability), and the systematic error (any uncorrected bias or offset of the measurements relative to national or other designated standards).

3. ISN'T IT POSSIBLE TO ACHIEVE A HIGH LEVEL OF ACCURACY BY USING NBS CALIBRATION SERVICES INSTEAD OF NBS MAP SERVICES?

Yes, but experience has disclosed that some users of NBS calibration services have had longstanding measurement problems that remained undiscovered until they participated in a MAP. It is certainly possible for a laboratory to achieve a high level of accuracy without using NBS MAP services if standards calibrated by NBS are used to assess the offset of the measurement process from the nationally accepted reference base and if rigorous measurement quality assurance procedures described elsewhere (see Part II of this guide) are used.

Occasionally participation in a MAP discloses that a laboratory is performing more accurate measurements than had been assumed. New MAP participants often find, however, that their measurement uncertainty is not as good as they had thought. Participation in a MAP, often improves the laboratory's precision or accuracy from initial values. Because the measurement assurance regimen requires that measurements be made on an on-going basis following consistent procedures, some facility may be acquired in following the measurement procedures that did not previously exist. In other cases, flaws in the measurement methods or environmental conditions had gone unnoticed when the laboratory relied only on NBS calibrations.

4. HOW DOES A MAP SERVICE DIFFER FROM A "ROUND-ROBIN" INTERCOMPARISON?

Round-robin** intercomparisons of standards are often used to reveal systematic errors and measurement inconsistences among laboratories,

* Defined in the glossary.

** Round-robin is defined in the glossary.

but a MAP is more than a round-robin intercomparison. In order to take full advantage of a MAP service the participant is expected to make measurements on a continuing basis, using an in-house check standard between the times that the transport standard is measured, to provide assurance that the measurement process has not gone out of control since the last measurement on the NBS transport standard.

5. HOW MANY MAP SERVICES DOES NBS CURRENTLY OFFER?

NBS currently (1984) offers eight MAP services in the following areas:

- Mass
- Gage Blocks (pilot program)
- DC Voltage (standard cells)
- Capacitance
- Resistance
- Electric Energy (watthour meters)
- Temperature (resistance thermometry)
- Laser Power and Energy

Other MAP services are being developed in areas such as microwave power and spectrophotometry. Chapter 5 provides additional details.

6. HOW MUCH DOES IT COST TO USE NBS MAP SERVICES?

The cost of the NBS MAP service depends on the service and how frequently it is used. The actual cost per year to the participant may average less than the currently advertised MAP service cost since it may not be necessary to use the service annually. Also, the organization may choose to join with others in a group arrangement (described in Chapter 6) to reduce the cost. The fees for NBS MAP services change from time to time; thus it is advisable to check with the point of contact listed in NBS Special Publication 250 to determine the current price*.

In addition to the NBS fee, participation in a MAP may involve costs associated with the purchase of additional equipment and/or additional staff time in the participating laboratory, particularly if the laboratory has not previously instituted quality control procedures in its measurement process.

7. IF I UTILIZE AN NBS MAP SERVICE, WHAT DO I HAVE TO DO BESIDES MAKE MEASUREMENTS ON THE TRANSPORT STANDARD?

It is important to recognize that NBS does not audit or regulate metrology laboratories as part of the MAP service. Whatever steps are taken by a laboratory participating in a MAP to improve its measurement process are taken voluntarily.

---

*Kieffer, L. J. ed. Calibration and Related Measurement Services of the National Bureau of Standards, Natl. Bur. Stand. (U.S.) Spec. Publ. 250; (new edition issued every two years). An Appendix to SP250 (current price list) is published by NBS every six months.

Although the transport standard calibrated by NBS is indispensible in the operation of a MAP, participation also requires making measurements on in-house check standards on a continuing basis to estimate the random error and to make sure that the measurement process remains in a state of statistical control*. Unless the participating laboratory has a measurement control program to monitor its own measurement process parameters, there is little point in participating in a MAP service.

MAP participants may also have to perform some data analysis.

8. DO I HAVE TO BE AN EXPERT STATISTICIAN TO BE ABLE TO PARTICIPATE IN A MAP?

No, all one needs is "statistical awareness," that is, an appreciation of the rationale for the statistical techniques. However, the more one knows about statistics, the better one will be able to interpret and utilize the results. The amount of statistical manipulation of data by the MAP participants varies among the existing NBS services from essentially none to a considerable amount. In general though, all that is required is that measurement instructions be followed and data reported in a specified format. The amount of data analysis done by NBS as compared with that done by the participants is negotiable. For many MAPs, the data can be analyzed on a programmable calculator. NBS can provide participants with tapes or listings of many of the programs. (See Chapter 5 for details.)

An individual with some knowledge of statistics will be able to understand most of the MAP methodology used by NBS in the data analysis. Someone with a more extensive knowledge of statistics will be able to appreciate all of the subtleties of the method and may be able to see ways to utilize the data more effectively. NBS staff are eager to have each participant succeed and will provide consulting help whenever necessary to explain the data analysis and methodology.

9. DOES A MAP SERVICE PROVIDE THE TRACEABILITY TO NATIONAL STANDARDS REQUIRED FOR COMPLIANCE WITH MILITARY SPECIFICATIONS OR REGULATORY DOCUMENTS?

Successful participation in a MAP provides excellent evidence of traceability to national standards. Users of MAP services receive a test report from NBS stating their measurement uncertainty.

NBS does not require traceability of anyone, nor does NBS have legal responsibility for determining whether or not a particular organization has adequately demonstrated traceability to national standards. This is the responsibility of auditors from the organization requiring traceability (e.g., the Defense Contracts Administrative Service, the Nuclear Regulatory Commission, etc.)

---

* "State of statistical control" is defined in the glossary.

Traceability to NBS has traditionally been achieved by obtaining an NBS calibration of customer-owned standards. Prior to the introduction of MAP services, auditors generally considered an organization to have met the requirements for traceability if documentation could be produced to show that its standards had been calibrated "traceable to NBS." When MAP services were first introduced, some auditors who were unfamiliar with the approach questioned the acceptability of the MAP reports as evidence of traceability, since the MAP participant's standards were not calibrated by NBS. The problem now seems to be disappearing as auditors come to appreciate that a MAP is usually a more effective kind of traceability than an NBS calibration. This is discussed in more detail in Chapter 4.

10. WHAT EXACTLY DO I GET FROM NBS WHEN I REQUEST A MAP SERVICE?

Typically, the customer receives from NBS (usually by air freight) one or more transport standards that have been carefully measured before leaving NBS. The standard is measured a prescribed number of times by the participant and returned along with the data to NBS. NBS remeasures the standard and then provides a test report stating the offset of the participant's measurement results from NBS and the associated uncertainty. Usually, NBS provides some or all of the data analysis. NBS also provides technical publications and/or oral guidance on theoretical considerations, measurement control techniques, and recommended practices for the various measurements. When a problem arises in the participating laboratory, NBS will also provide (within reasonable limits) consultation to uncover and correct the problem. (If the customer does not already employ measurement quality control practices, NBS will provide material explaining how to institute such practices, before sending the transport standard.)

11. WHAT IS A REGIONAL OR GROUP MAP?

This new approach to disseminating MAP Services is described in detail in Chapter 6. Briefly, a regional or group MAP is a MAP wherein cooperating laboratories interact with NBS as a group. Generally one laboratory agrees to serve as the "pivot" laboratory, providing the principal point of contact with NBS. The out-of-pocket cost to the participants in a group MAP is reduced by sharing the cost of the transport standard from NBS. Faster resolution of measurement problems and other benefits may also result from group participation.

Those considering MAP participation on a regional basis are encouraged to call or write to the chairman of the Measurement Assurance Committee of the National Conference Standards Laboratories (See Chapter 8). The prospective participant can then be put in touch with other similar laboratories who have participated in group or regional MAPs, and learn of their experiences.

12. HOW DO MAP SERVICES RELATE TO LABORATORY ACCREDITATION?

NBS does not presently accredit calibration laboratories for all types of calibrations, although some limited scope calibration accreditation programs are receiving consideration under the auspices of the National Voluntary Laboratory Accreditation Program, NVLAP*. Ideally, accreditation should be based on a laboratory's ability to demonstrate that its measurements have uncertainties relative to national standards less than some specified limit. Successful participation in a MAP can provide important evidence of competence required for laboratory accreditation by any organization that chooses to accredit laboratories.

13. WHAT SPECIAL PRECAUTIONS NEED BE TAKEN WHEN MAKING MEASUREMENTS ON THE NBS MAP STANDARD?

The MAP service is designed to assess the quality of the laboratory's calibration process, hence it is essential that the measurements on the transport standard reflect the normal operating conditions of the laboratory. Because future assignments of values to items calibrated by the laboratory will depend on these measurements, the laboratory's measurement system must be operating in a state of statistical control when the comparison with the transport standard takes place. To ensure process control operational and statistical tests or checks should be included in the measurement scheme when the transport standard is measured by the participant. Such checks may or may not be part of the instructions issued by NBS. Upon request, NBS can provide guidance on suitable checks where such checks are not explicitly included in the instructions.

Strictly speaking, the offset identified by the exchange with NBS applies only to laboratory conditions that are identical to test conditions. For example, optical systems with visual eyepieces are operator-dependent, requiring separate tests for each operator, and resulting in individual offsets or corrections for each operator. Extension of the uncertainty statement to varied laboratory conditions is valid only insofar as the error estimate has been structured to include these variations.

14. MUST A MAP BE OPERATED AT STATE-OF-THE-ART ACCURACY?

No, one must distinguish between the MAP concept and NBS MAP services. NBS MAP services are generally intended to be at state-of-the-art accuracy, but the MAP concept can be applied at any level of accuracy.

---

* For more information on the NVLAP accreditation program, write to the Office of Laboratory Accreditation at NBS.

15. CAN LABORATORIES OUTSIDE THE U.S. PARTICIPATE IN MAPS?

Under very special circumstances, this may be possible, but requests for such participation must be reviewed on a case-by-case basis. MAP services are not normally provided to non-U.S. requestors. Technical constraints on the long-distance transporting of standards may also limit participation from outside the U.S. A foreign laboratory should transmit its request for NBS MAP services to the NBS Office of International Relations along with an explanation of why measurement services available in its own country are not adequate. To expedite the decision on the request, a letter from the national standards laboratory or embassy of the country should accompany the request. This letter should indicate that the government of the requesting country has no objection to NBS providing such service.

When NBS can grant the request, the cost to the foreign participant will exceed that to U.S. participants due to the additional costs of communication and shipping.

16. HOW CAN THE COST OF A MAP SERVICE WHICH MAY EXCEED THAT OF AN NBS CALIBRATION BE JUSTIFIED TO COST-CONSCIOUS MANAGEMENT?

The nature of the justification will vary depending on the mission of the laboratory. The MAP service provides an unambiguous mechanism for proving the competence of a laboratory in performing accurate measurements relative to national standards.

To quantify the benefits of a MAP, the lab manager must ask, "What is the economic penalty associated with having measurements of inadequate accuracy and/or unknown uncertainty?" Inadequate measurement capability in industry often leads to "good" products being scrapped or submitted for rework and "bad" products being accepted. It may lead to costly disputes between a firm and its suppliers or its customers. The resulting economic penalties can often be estimated. One can also estimate the costs associated with having to overdesign a product to meet a tight specification because of the inability to accurately measure its actual properties.

While a MAP service generally costs more than the corresponding calibration service on a one-time basis, one should consider the cost differential from a long term perspective. After participating in a MAP for a period of time, most participants find that they can extend the intervals between transfers from NBS so that the NBS MAP service is used less frequently than the corresponding calibration service. Thus, the use of MAP services may be more expensive in the short term, but less expensive in the long term.

17. HOW WOULD AN AUDITOR CHECK A LABORATORY UTILIZING NBS MAP SERVICES TO ENSURE THAT ADEQUATE TRACEABILITY WAS BEING MAINTAINED?

The relationship between MAPs and traceability is discussed in Chapter 4. The procedure for auditing a laboratory utilizing NBS

MAP services would not differ significantly from that used now to audit laboratories relying on conventional calibration hierarchies tied to NBS calibrations. The auditor would still ascertain that an adequate overall quality plan is available and that the procedures actually used are those documented in the plan. The laboratory can make the MAP test reports from NBS and the MAP data available for inspection by the auditor.

18. DOES NBS DICTATE TO THE MAP PARTICIPANTS WHAT PROCEDURES AND EQUIPMENT MUST BE USED?

NBS provides recommended measurement procedures to a greater or lesser extent for every MAP service. For some MAPs the participant has considerable latitude in choosing which particular apparatus and techniques are to be used. In other cases, NBS strongly advises that certain procedures be followed. If you are interested in a particular MAP service, discuss your concerns with the NBS point of contact listed in Special Publication 250. More often than not NBS will have enough flexibility in the program to accommodate your needs.

19. TO WHAT EXTENT DOES NBS RESPECT THE PRIVACY OF THE PARTICIPANT? TO WHAT EXTENT WILL NBS RESPECT THE WISHES OF THOSE PARTICIPANTS WHO WANT TO KEEP THEIR MAP DATA PROPRIETARY?

NBS calibration reports and MAP test reports are considered to be the exclusive property of the customer, and NBS does not release these reports to third parties (including regulatory agencies) without the express permission of the customers.

As part of the requirements of Defense Department contracts or regulatory compliance, auditors will presumably want to see the MAP test reports just as they now wish to see NBS calibration test reports.

The question is sometimes asked of NBS, "What would happen if some third party were to request an NBS MAP test report under the Freedom of Information Act?" (A MAP test report would be no different from a calibration test report in this regard.) A general answer to this question cannot be given. As indicated above, NBS does consider test reports to be proprietary, but the criteria for releasing information under the Freedom of Information Act are complex, and those who wish more information on this topic should contact the NBS Legal Office.

In the case of a group or regional MAP, a few participants have been hesitant about discussing MAP data with other group members initially, and NBS staff have respected their wishes. After the group MAP is underway and each participating laboratory gains confidence that its measurements are under control, the group participants often develop a spirit of camaraderie and generally share data and experiences, helping each other continue to improve the group's performance.

3. The Philosophy of a MAP*

3.1 Introduction

Often, action taken to maintain our health, safety or the quality of our environment is based on a single measurement. It is important, therefore, that the errors of such measurements be small enough so that the actions taken are only negligibly affected by these errors. We realize this necessity on a personal basis when we consider medical measurements, or measurements of our exposure to radioactivity. It is also obvious that the "shadow of doubt" surrounding the measurements should be suitably small in any government regulatory action or measurements involved in legal actions. This is no less true for all other measurements in science and industry; even though legal action may not be involved, the validity of scientific inference, and the effectiveness of process control on the quality of production depends on adequate measurements.

3.2 Measurement Credibility

Consider what might happen when a measurement becomes the subject of a legal or scientific controversy--when its credibility as scientific evidence is in question. As with other items of evidence, the "shadow of doubt" (the size of the uncertainty) associated with the measurement has to be determined. The measurement must be able to stand up under "cross-examination." Circumstantial evidence such as the brand name or high cost of the instrument or even its recent calibration by an "NBS-traceable" laboratory may not be sufficient evidence of good measurement, particularly when the conditions of field measurement are hostile. One needs evidence of measurement quality that will "stand up in court."

Regarding measurement in the context of a cross-examination brings a number of issues into focus. The amount of required evidence depends on the use of the measurement. A once-a-month check on measurement performance may be adequate for some measurements; for others it may be necessary to measure reference standards both before and after the measurement of interest. For any important measurement, a statement should be developed as to what is "good enough" in measurement, i.e., the uncertainty which can be tolerated, because errors smaller than that threshold contribute only negligibly to the correctness of the decision one makes using the measurement.

* Excerpted with editing and additions from "Measurement Assurance" by J. M. Cameron (Bibliography item number 1) and unpublished material by Cameron.

3.3 Allowable Limits of Measurement Error

How does one determine whether particular measurements are "good enough" for their intended use? What is "good enough"? There are a number of cases where physiological or other constraints provide a criterion. In treating cancer with cobalt radiation, too much radiation destroys healthy tissue surrounding the tumor, resulting in adverse side effects to the patient. If too little radiation is applied, the tumor is not destroyed and the malignancy continues. The medical profession has established tolerance limits on radiation dosages required in order to achieve good cure rates for particular kinds of tumors. In this instance, the definition of a "good measurement" has a firm physiological basis in terms of the error permitted in exposure to cobalt radiation.

In nuclear materials control the allowable error is a function of the amount of material which would pose a hazard if diverted (e.g., enough to build a bomb). In industrial production or commercial transactions, the error limit is determined by a balance between the cost of better measurement and the possible economic loss from poorer measurement.

By whatever path such requirements are arrived at, let us begin with the assumption that the allowable or maximum permissible error should not be outside the interval $\pm$ a, where a is some stated uncertainty or tolerance. In the more general case the uncertainty may be stated as +a, -b relative to the quantity being measured. Our problem is one of deciding whether the uncertainty of a single measurement is wholly contained in an interval of that size. We, therefore, need a means of assigning an uncertainty to a single isolated measurement.

In order to give operational meaning to the term "uncertainty," we need a perspective--a physical and mathematical model--from which to view measurement.

3.4 Reference Base to Which Measurements Must be Related

In the "cross-examination" to determine the credibility of a measurement, a logical first step would be to ascertain what approach or approaches to measurement the contending parties would view as acceptable. In other words, it is necessary to establish what a reasonable, prudent, technically-competent person would do in measuring the quantity in question. If agreement cannot be reached on this point, it is not possible to determine whether or not the measurement in question was adequate (of acceptable uncertainty).

The term "true value" is sometimes used to refer to the correct or actual value of the quantity being measured. Since instruments are never perfect, measuring environments change, etc., no one can know the true value of any measured quantity; however, one can usually make measurements that approximate the true value sufficiently closely to meet his or her objectives.

Through international agreements, reference bases and standards are established for nearly all measurements of interest, such as temperature fixed points for thermometry and certain numbers of wavelengths of light to establish the unit for dimensional measurements. The mass of the prototype kilogram standard kept by the International Bureau of Weights and Measures (BIPM) near Paris is known to have a mass with a true value of exactly one kilogram because this artifact is defined to be exactly one kilogram by mutual agreement of all nations.

Because no one can determine the true value for most measurements, the term "consensus" or generally-accepted value is used to refer to the value that would be obtained from a comparison of an unknown to these agreed-to standards using accepted practices. The consensus or generally-accepted value has a particularly simple meaning for measurements of such quantities as mass, voltage, resistance, temperature, etc. One may require, for example, that uncertainties be expressed relative to the standards as maintained by a local laboratory or, when appropriate, to the national standards as maintained by NBS. In other cases, nationally accepted artifacts, Standard Reference Materials or, in some cases, a particular measurement process may constitute a reference base. In any event, the selected and agreed-upon reference base must be realizable in the real world.

3.5 Properties of Measurement Processes

For measurement assurance purposes, it is useful to regard measurements as the "output" of a process, analogous to an industrial production process. Two characteristics of measurement processes should be noted. Firstly, repeated measurements of the same quantity by the same measurement process will differ slightly (assuming the process has sufficient resolution), and, secondly, the averages of long series of measurements by two different processes will generally differ somewhat. The conditions must be defined under which a "repetition" of the measurement would be made, analogous to defining the conditions of manufacture in an industrial process.

3.5.1 Disagreement Among the Measurements

One's early experience with measurements, e.g., using a ruler in school, usually involves a coarse enough interval so that either successive measurements of the same quantity agree or the variation from item to item of material is large compared to the measurement error. As requirements for accuracy and precision increase, however, measurement is characterized by the fact that repeated measurements of the same quantity disagree to a significant extent (e.g., measuring gage blocks using a high precision comparator). This disagreement may be due to environmental factors such as temperature, pressure, and humidity, or to changes in procedure, operator techniques or instruments. The variation may also be due to shortcomings in the physical model related to nonlinear response, hysteresis, interference from other phenomena, etc.

3.5.2 Measurement as a Production Process

When a sequence of measurements is made on the same object the analogy between a measurement process and a production process becomes more apparent. Figure 1 is a plot of repeated determinations of the mass of a nominal 10 gram weight made during the period 1963 to 1965. The dotted line indicates the mean of the measurements; the solid lines are control limits within which the measurements are expected to lie. Part II of this guide explains in detail how these limits are established. All

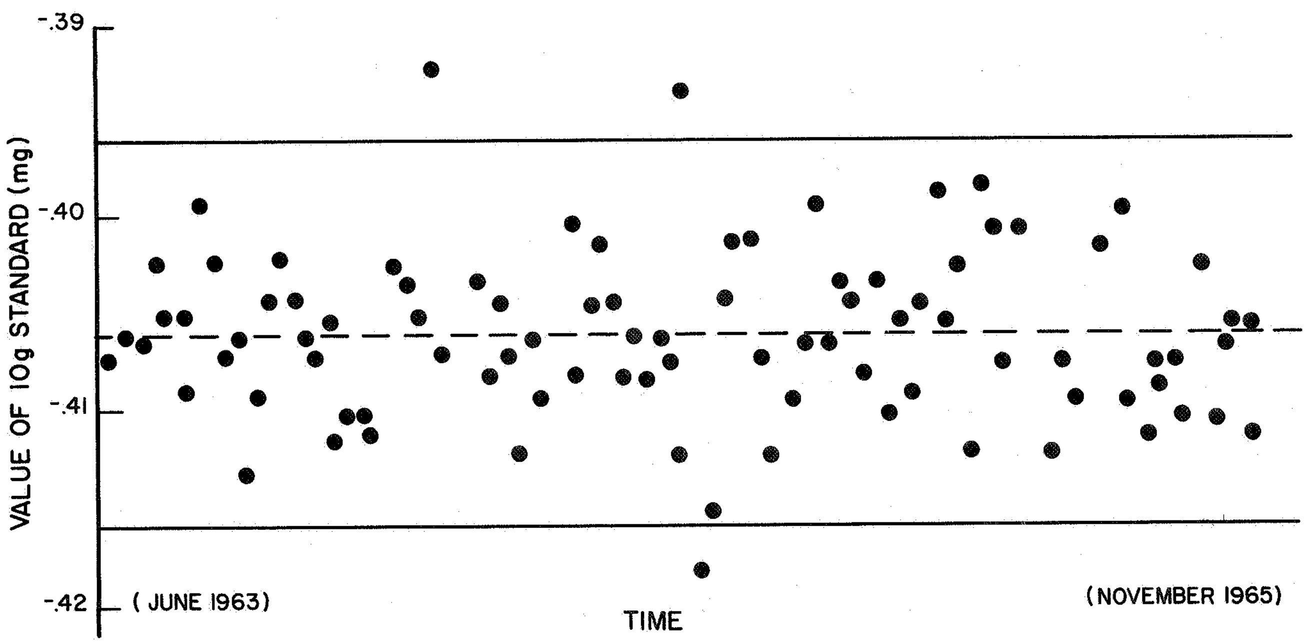

FIGURE 1: DEVIATION OF A NOMINAL 10 GRAM MASS STANDARD FROM ITS NOMINAL VALUE.

the measurements recorded in Figure 1 are the result of applying nominally the same procedures on successive occasions. The lack of perfect agreement among the results is due to variations in the execution of the process of measurement, imperfections in the instruments, and in controlling or correcting for ambient conditions. However, the set of measurements of Figure 1 appears to be measuring the same underlying quantity, i.e., the measurements appear to have the same limiting mean (long term average). Further, the variation of the points about the central line (the limiting mean)* appears to be random. If these conditions are present, the process may be said to be in "a state of statistical control."

If the measurements are the result of a process in a state of statistical control, then by determining and describing the parameters of that process, the behavior of further measurements can be predicted and an uncertainty can be assigned to an arbitrarily selected measurement.

3.5.3 Model of the Measurement Process

To characterize a measurement process, one strives to develop a model to explain the observations. It is helpful to think of the complete model as consisting of a physical part and a mathematical (statistical) part. (Mathematical expressions may, of course, be used to describe the physical part.) The physical part consists of all factors known to affect the process and a description of their influence. For example, a model for a particular process might include the knowledge that the measured value of a standard increases linearly with temperature.

The statistical portion of the model is a description of the variability arising from all causes not explicitly accounted for. In the above example, there may be errors in the determination of the actual temperature of the standard even though the temperature dependence is known. Any variability between the actual temperature and the measured temperature of the standard will, of course, lead to variability in values assigned to unknowns being calibrated relative to the standard.

The development of a complete model for a measurement process is typically an iterative process. One begins by considering obvious sources of variability and bias and attempting to quantify them through experiment or other means. As more and more factors are successfully accounted for, the model improves and the unexplained sources of error decrease.

The list that follows gives examples of questions that might be asked during the development of a model of a measurement process:

> Within what limits would an additional measurement by the same instrument agree when measuring some stable quantity?

---

* Limiting mean is defined in the glossary.

Would the agreement be poorer if the time interval between repetitions were increased?

What if different instruments from the same manufacturer were used?

If two or more types (or manufacturers) of instruments were used, how much disagreement would be expected?

To these can be added questions related to the conduct of the measurement. For example:

What effect does geometry (orientation, etc.) have on the measurement?

What about environmental conditions, temperature, moisture, etc.?

Is the result dependent on the procedure used?

Do data taken by different operators show persistent differences in values?

Are there biases or differences among nominally identical instruments due to reference standards or calibrations?

These questions serve to define the measurement process--the process whose "output" is to be characterized.

As an example, a sequence of measurements was made using two sound level meters to measure a noise level of nominally 90 dB referenced to 20 Pa*. The sound was generated by a loudspeaker fed a signal from a broadband noise source. On 16 different days, measurements were made outdoors and over grass with the loudspeaker in the same orientation and location relative to a building 2 m behind the loudspeaker. The sound level meter was always the same distance (10 m) from the loudspeaker and on a line perpendicular to the face of the loudspeaker. There were no other reflecting surfaces or obstacles within 50 m. No measurements were made in the rain or in winds exceeding a few km/hr. The averages for each day's measurements are shown in Figure 2.

To illustrate variation within a given day's measurements, individual measurements made on one meter at a different orientation are plotted in Figure 3.

The data from the sound level meter example (Figures 2 and 3) are more complex to analyze than the mass data (in Figure 1). It appears that the mass data are from a measurement process characterized by a stable mean value, with a certain amount of random scatter about that mean value. In the sound level example, there appear to be shifts in the mean from day to day in addition to the variation observed within any particular day's readings. The statistical model for the mass case is less complex. For example, the data suggest that future observations will be normally distributed about a stable mean value; the mean value and the scatter for future measurements can be predicted from the data already gathered.

---

* Magrab, Edward B., Environmental Effects on Microphones and Type II Sound Level Meters. Natl. Bur. Stand. (U.S.) Tech. Note 931; 1976 October.

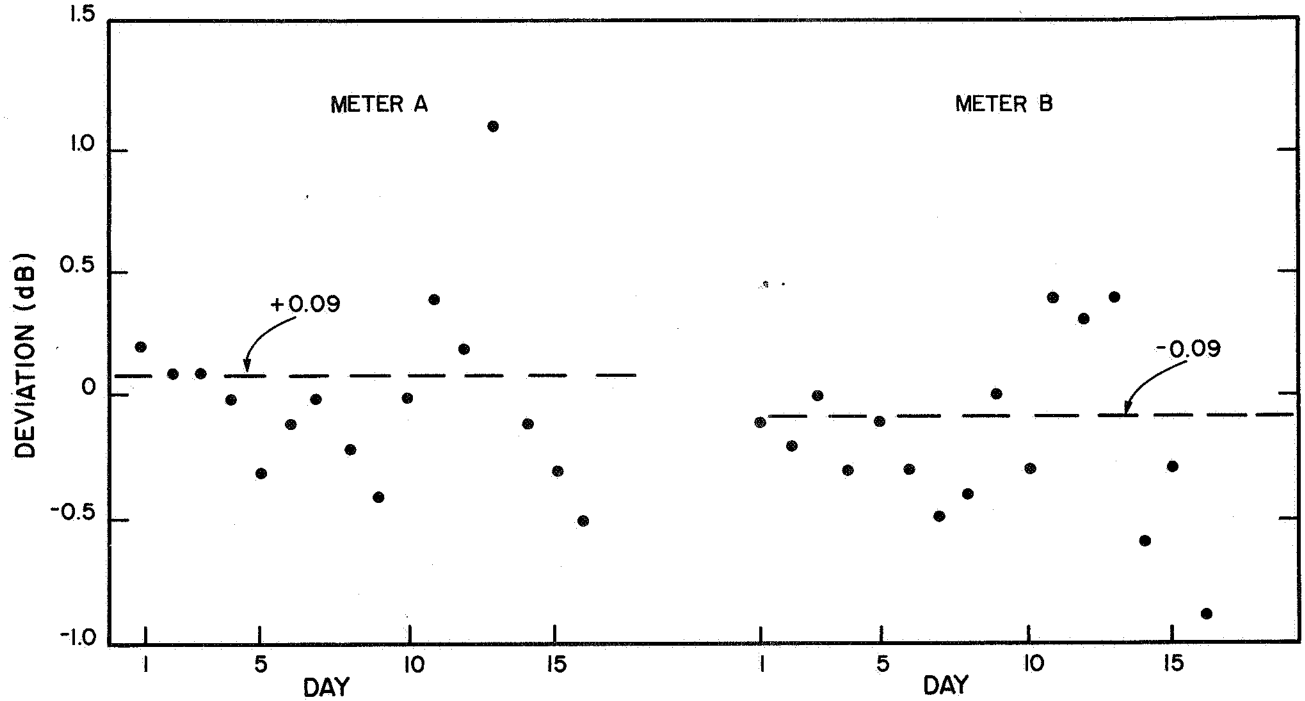

FIGURE 2. DAY-TO-DAY VARIATION IN AVERAGE METER READINGS.

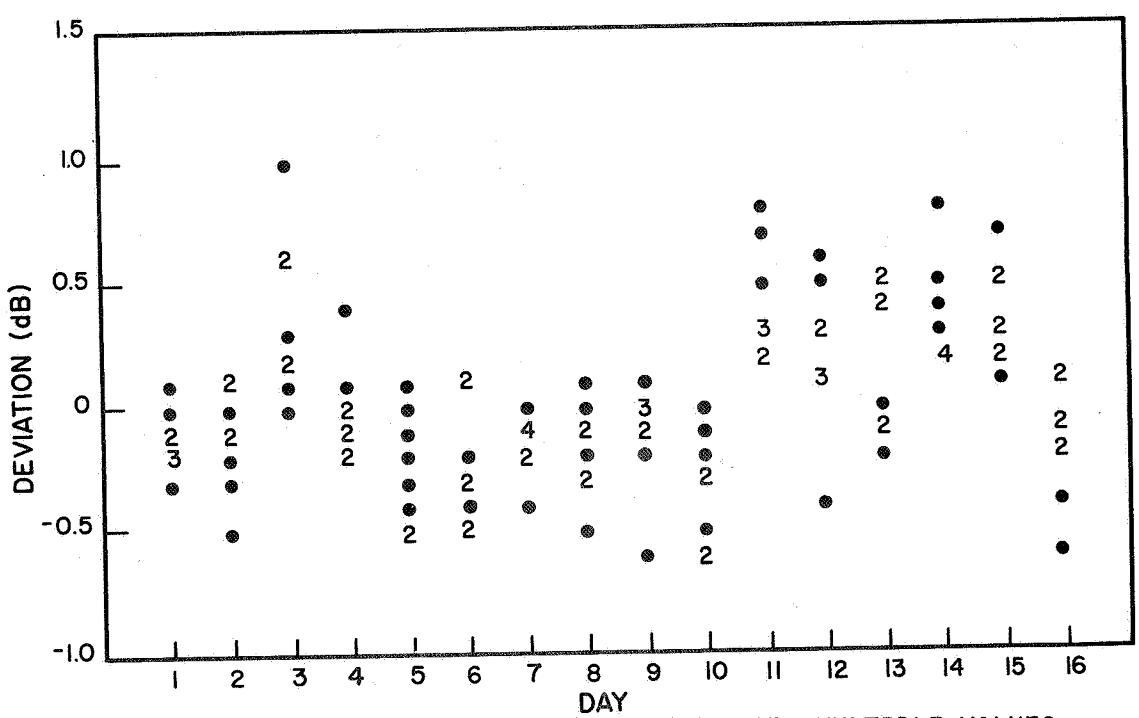

FIGURE 3. DAY-TO-DAY VARIATION IN A METER READING WITH MULTIPLE VALUES PER DAY. (COINCIDENT POINTS INDICATED BY NUMBERS.)

The statistical model for the sound level example would probably have to include two components of variation--one to account for the within-day variation and a second component to account for the variation in the mean of the measurements from one day to the next. More data would probably have to be collected in order to characterize these measurements and quantify the uncertainty.

In modeling a measurement process one examines data such as those in the examples above, postulates a model, tests the model for consistency with the data using statistical tests (described in Part II of this guide and revises the model as required in an iterative fashion until satisfactory agreement is achieved between the model and the observed data.

3.6 Offset or Systematic Error

Despite attempts to quantify the factors influencing a measurement or to eliminate their effects by the use of corrections, there usually remain systematic differences (biases) between results made under different conditions. When different measurement processes are used to measure the same quantity, for example, systematic differences are to be expected. In case of controversy, which measured value is correct? The answer involves defining a reference base to which offsets or systematic errors can be referenced.

For mass, length, voltage, and many other quantities for which NBS maintains national standards, the offset of one's measurement process can be determined directly through the use of NBS calibration services or MAP services, or indirectly through the use of services from a high level commercial calibration laboratory that uses NBS services. For many types of chemical or materials properties measurements it is more appropriate to use NBS Standard Reference Materials to assess offset.

For those measurements where NBS does not maintain national standards or provide services, it is still necessary to estimate the possible offset of one laboratory's measurements from those of others. Often this is accomplished through intercomparisons among several laboratories making the measurements. A stable item or group of items is circulated among the participating laboratories and measured by each. The group mean (after obvious outliers are discarded) is then adopted as the reference base, and offsets of individual laboratories can be referred to this reference base.

3.7 Uncertainty

Part II of this guide describes in detail how an uncertainty statement for a measurement process is constructed. This section is limited to a qualitative treatment of the subject of uncertainty.

In assessing the uncertainty of a measurement process, questions such as the following should be asked:

- What errors can arise from unknown departures of environmental or operating conditions from nominal values?

- What systematic errors could result from departures from the assumed physical model due to unaccounted-for nonlinearity, hysteresis, time dependent effects, etc.?

- What is the possible offset of the process from the reference base?

- Does a state of statistical control exist so that historical data can be used for setting bounds to the effects of random error?

In assessing the uncertainty of measurements, all significant factors contributing to the offset and the random error must be considered. The analogy of measurement with industrial production processes is useful in understanding this point. If the variability of a product were reduced to a negligible level, all end-product items would be at the process average, which would have a fixed relationship to the specifications. The limiting mean of a measurement process has a similar relationship to the true value for the measurements as the process average for a manufacturing process does to the specifications. Any uncorrected difference between the limiting mean and the reference base is known as the systematic error of the measurements. Where the random error is negligible, the uncertainty consists entirely of the systematic error. If random errors are present, the uncertainty will be further increased by a suitable measure of the possible magnitude of such random errors. The offset of the process is certainly not changed by the presence of random error, and the limits for the random error are independent of the offset of the process limiting mean from the reference base. The uncertainty is the sum of these two distinct components, the offset and the limit of random error.

3.8. The Concept of a Repetition and the Check Standard

Two requirements for demonstrating the validity of a measurement process are: i) predictability that the variability or scatter will remain at the same level, and ii) evidence that the process will not drift or shift abruptly from its established values (or that if there is drift, it can be predicted and corrected for).

In cases where the measurement can be repeated, one can determine random errors by remeasuring at a later time sufficiently far removed to guarantee independence of the measurements. For many measurement processes, time, availability of personnel and equipment, or other constraints often limit the determinations on any given item or unknown to a single measurement. How is it possible in these instances to make a statement regarding the scatter of the results that would have been expected had there been multiple measurements of any particular unknown?

To characterize the random error of a measurement process, some redundancy needs to be built into the measurement scheme. This redundancy is usually obtained through repeated measurements on a stable item (or items) called a check standard*. When there exists a sufficiently large historical database of check standard measurements that are similar to the measurements being made on unknowns, limits can be established within which the next measurement on the check standard would be expected to lie, and it can be assumed that the scatter of multiple measurements on

* Check standard is defined in the glossary.

an unknown would be comparable to that for the check standard. For these limits to be valid, the measurements on the check standard must be independent (uncorrelated). In other words, the check standard measurements must encompass a sufficiently broad range of environmental and operating conditions to include all of the random effects to which the process is subject. When this situation exists, one can legitimately claim that the random error components of the measurements on the check standard and the measurements on the unknown will be comparable in magnitude.

Thus suitable data for estimating random errors can be obtained by incorporating an appropriate check standard into routine measurement procedures, provided the check standard is subject to the same variability to which the "unknown" is subject. The statistical procedures for expressing the results will depend on the structure of the data but cannot overcome deficiencies in the extent to which the check standard measurements simulate measurements in the unknowns.

In mass calibrations at NBS, a check standard is measured in conjunction with the unknowns. In this simple case the check standard is treated in exactly the same way as the unknowns so that the behavior of the process with respect to the check standard is transferable to the unknowns. A sequence of measurements of the same check standard is generated for an extended time period.

The data on the check standard provide the basis for quantifying the random error of the measurement process, assumed to be transferable to the unknowns. One is saying, in effect, if the "unknown" were measured again and again, a sequence of values such as those for the check standard would have been obtained. Whether the single value obtained is above or below the limiting mean cannot be determined, but it is fairly certain the single value would not differ from the limiting mean by more than the bounds to the scatter of the values on the check standard.

3.9 Measurement Process Control

An uncertainty statement is only valid when the measurement process is in a state of statistical control. Once an out-of-control condition occurs, predictability is lost.

In routine mass calibration at NBS, a check standard is included with each set of weighings of unknowns, and process control is monitored by monitoring the value obtained for the check standard. The random error is determined from an analysis of the check standard data using "least squares" techniques described in Part II of this guide. Control charts in mass calibration have been maintained at NBS since 1963. In the calibration of gage blocks, similar process control has been maintained since 1972 on both the NBS interferometric process (by which the lengths of the NBS master gage blocks are assigned) and on the NBS comparator process (by which length values are transferred to customer gage blocks.) Thus, incorporation of check standards into a routine measurement process achieves two objectives; monitoring process control and assigning random error limits to measurements on unknowns. The sections of the bibliography dealing with mass and gage blocks list references that provide more details on these particular measurement control techniques.

When MAP techniques are applied in situations outside a traditional standards laboratory environment, a variety of constraints usually necessitates some flexibility and innovation to develop and implement the program. The uranium hexafluoride (UF_6) cylinder program for nuclear safeguards* is an example of techniques developed by NBS to provide a direct method for determining the offset of a practical mass measurement process from NBS mass measurements. The reference given in the footnote describes the particular constraints that had to be dealt with in developing this program.

Procedures have been established** to monitor the output of firms calibrating personnel dosimeters. In this case, a table was prepared of allowable limits of uncertainty based on physiological considerations. A model of the measurement process was determined in an initial study. Check standard data is monitored routinely to confirm that the process is in control at the specified levels. These "consistency" or "in control" criteria replace the round-robin approach previously used. In both the dosimetry and the UF_6 cylinder examples, there is firm evidence that the MAP methods improved the quality of the measurements that were subjected to these controls.

Section 8.5 of this publication provides additional examples of the application of these techniques to specific measurement requirements.

---

* Doher, L. W.; Pontius, P.; Whetstone, J. A New Approach for Safeguarding Enriched Uranium Hexafluoride Bulk Transfers. Nuclear Safeguards Technology 2; 1978; IAEA-SM-231/68. (Published by the International Atomic Energy Agency, Vienna, Austria, 1979.)

** Criteria for Testing Personnel Dosimetry Performance. Health Physics Society Standard; Working Group 1.4; 1981 June.

4. The Relationship of MAPs to Traceability to National Standards

NBS receives frequent inquiries concerning the term "traceability" and phrases such as "traceable to standards maintained by NBS." Requirements for traceability to NBS are frequently found in government procurement contracts (in particular, Department of Defense contracts) and in government regulations (e.g. those of the Nuclear Regulatory Commission and the Food and Drug Administration). Traceability requirements are established and enforced by the agencies requiring traceability of their contractors or those being regulated. NBS has no legal authority to determine whether or not a particular party has achieved adequate traceability. Moreover, agencies requiring traceability to NBS may not always agree on what constitutes adequate traceability.

One definition of traceability is given in Military Standard 45662 "Calibration Systems Requirements": "The ability to relate individual measurement results to national standards or nationally accepted measurement systems through an unbroken chain of comparisons." The traditional realization of this and most other definitions of traceability is through a hierarchical calibration system. A hierarchical calibration system is one in which a "primary" or high level laboratory calibrates the lower accuracy standards of a "secondary" or intermediate level laboratory. The intermediate laboratory then calibrates field instruments of still lower accuracy. Some hierarchical calibration systems operate with several steps between the highest level laboratory and the field measurements.

The use of a hierarchical calibration system for traceability provides a necessary condition for accurate measurements, but unless it is coupled with an appropriate degree of measurement control at every step of the hierarchy, it is not a sufficient condition. A hierarchical system can be very effective if adequate internal measurement controls are incorporated to ensure that the transfers between the levels within the hierarchy are performed without introducing unacceptable errors. Agencies requiring traceability to NBS generally would agree that the use of an NBS MAP service is evidence of traceability.

There are a variety of ways that traceability to NBS can be achieved. Techniques commonly used to establish and maintain traceability to NBS are:

1. Calibration of standards or instruments by NBS. This is the most traditional method for the realization of traceability to NBS; however, NBS calibrations must be supplemented by additional quality assurance procedures within the laboratory using such calibrated standards.

   Annually, NBS calibrates several thousand items for approximately 1000 different organizations (industry, other Federal agencies, state and local governments, universities, etc.)

2. Measurement Assurance Program Services. NBS has a number of measurement assurance services currently available (see discussion in Chapter 5). NBS MAP Services provide measurement quality control

procedures for calibration laboratories. These quality control procedures are combined to quantify the total uncertainty of measurements produced by the participating laboratory.

3. Standard Reference Materials (SRMs). SRMs prepared and sold by NBS support a wide variety of measurement accuracy requirements. These well-characterized reference materials are used in the chemical, biological, medical, and environmental fields. A surprising number of physical measurements can be supported by SRMs. Linewidth standards for the integrated circuit industry, for example, are now available in the form of an SRM. An SRM can be used to determine offset of a measurement process as well as become part of the check standard system to quantify random error.

   NBS currently has approximately 1000 different SRMs available.*

4. Time and Frequency Information. NBS disseminates time and frequency information over radio stations WWV, WWVB, and WWVH, television signals, and an experimental satellite service.** One or more of these NBS services together with appropriate internal quality assurance provisions will provide consistency with national time and frequency standards.

In addition to the use of the terms "primary laboratory" and "secondary laboratory" in describing the position of a laboratory in a hierarchical calibration system, the terms "primary" and "secondary" are also used in reference to traceability. If an artifact is directly calibrated by NBS or if an NBS SRM is purchased, evidence of primary (sometimes called "direct") traceability to NBS is said to exist. If a calibration is obtained from a laboratory whose standards are in turn calibrated by NBS, this constitutes secondary (sometimes called "indirect") traceability. Similarly, secondary traceability might be claimed by someone using a reference material purchased from an organization that used NBS SRMs to verify its measurement process on a regular basis.

In a hierarchical calibration system, the "primary" laboratory, of course, must quantify the total uncertainty of its measurements relative to NBS, or else the "secondary" laboratory cannot quantify its measurement uncertainty. By making repetitive measurements on its check standards, the secondary laboratory can establish bounds on the random error of its calibration process, but no matter how careful or conscientious it may be, it can only quantify its bias relative to national standards by means of the uncertainty statement accompanying the standards calibrated by the primary laboratory.

Some regard the ability to claim primary or direct traceability to NBS as a "status symbol." NBS, however, does not have the personnel to calibrate every reference artifact used in the U.S. and must encourage the development of capabilities both in the private sector and in government for providing secondary traceability. Calibration laboratories participating in MAPs can provide high quality secondary traceability services to others.

---

* Catalog of NBS Standard Reference Materials, Natl. Bur. Stand. (U.S.) Spec. Publ. 260 (new edition issued every two years).

** NBS Time and Frequency Dissemination Services, Natl. Bur. Stand. (U.S.) Spec. Publ. 432 (updated periodically).

5. NBS MAP Services

In this chapter each of the NBS MAP services currently available is described. References for each service can be found in the Bibliography under the heading "Specific Measurement Assurance Programs."

5.1 Mass (SP 250 No. 1.1B)

This service is most appropriate for primary calibration laboratories. Unlike most other NBS MAP services, the Mass MAP service does not involve the use of an NBS-owned transport standard that is shipped to the participants for measurement. The transfer standards in the mass MAP are a set of mass standards owned by the participant and sent to NBS for calibration. These standards are referred to as the "starting standards." In addition to the starting standards, the mass MAP participant must also furnish a set of much smaller weights called "sensitivity weights."

The choice of both the starting standards and the sensitivity weights will depend on the particular mass range of interest to the participant; NBS staff can provide advice regarding suitable starting standards and sensitivity weights for a particular range of mass weighings. In addition to the starting standards and the sensitivity weights, the participating facility should have a working set of weights known as the "test set" and a set of weights to be used as check standards. This set usually consists of weights in the range 1 g to 1 Kg.

This service, like other NBS MAP services, samples the participant's measurement process and establishes its uncertainty. Once the participant has become well-established in the Mass MAP, two options are possible:

1. NBS personnel do all of the data analysis and record keeping for the participant and provide periodic reports on the uncertainty of the participant's mass measurements.

2. NBS provides the participant with the methods and computer codes; the participant keeps all records and calculates uncertainties.

The implementation of the Mass Measurement Assurance Program in its most complete form typically proceeds in four distinct phases which may be abbreviated somewhat if the participant already has a suitable mass measurement quality control system:

Phase I

Participation in a Measurement Assurance Program for mass is generally initiated by a discussion at the managerial level between the appropriate NBS staff and the participant. This is sometimes followed by a several-day visit to NBS by the supervisor of the group that will be directly involved. Each new participant also completes a questionnaire on equipment and facilities. Normally a coordinator from NBS is named for each new participating laboratory. The NBS coordinator will become familiar not only with the weighing equipment (make and model) and available mass standards, but with present procedures and objectives of the participant.

The MAP participant receives a written description of the NBS process, methods and procedures to be used, an introduction to the interpretation of results, and information on the use of these results in measurement decisions. At the participating laboratory, the suitability of the weighing equipment is verified, the "starting standards" selected or procured, and if the procedures are entirely new, operators are trained. Throughout this phase, the NBS coordinator is available for consultation by telephone, letter or visit.

The starting standards and sensitivity weights are sent to NBS for calibration. Because the starting standards are to the participant as the reference kilograms are to the NBS process, any error in the value of the starting standards will show up as a systematic error in all future calibrations done by the participant. Therefore, the values for the starting standards are determined several times over a period of months at NBS. If the starting standards have a prior NBS calibration history, those data are reviewed, and if satisfactory, they are considered, along with the data from the more recent determinations, in arriving at assigned values for the starting standards.

The NBS coordinator recommends a weighing design to be used for calibrating the test weight. (The weighing design prescribes the set of observations for intercomparing the test weights with known weights.) The coordinator also supplies data sheets that are used throughout the first three phases of the program for recording data taken using the design. In the bibliography listings under the mass heading, reference 25 describes weighing designs for calibrating weight sets of various denominations.

During the first phase the NBS coordinator's objective is to make sure that the new participant is familiar with good laboratory practice for high precision weighing. If the participating laboratory has an established mass measurement capability and an existing quality control procedure for mass measurements, Phase I is abbreviated considerably.

Phase II

The starting standards and sensitivity weights are returned to the participant. Measurements are made over a period of time by the participant to verify that a state of statistical control exists. Following the prescribed procedures, the laboratory performs three or more independent calibrations of the test weight set using the starting standards and aforementioned weighing design.

The data sheets are sent to the NBS coordinator for review, comments, and processing. This is done sequentially (each experiment analyzed before the next one begins) for several reasons. If there are unanticipated problems, or the procedure has not been followed exactly, more measurements may be required. After three or more successful calibrations in the user's facility, the NBS Coordinator analyzes the data to determine the values of the check standards.

Phase III

A comprehensive report is issued by NBS at this point, which contains a review of the actions and decisions in each of the three phases, control charts for the check standards to be used in the participant's facility, and a comparison of the values assigned to the starting standards by NBS and by the participant. It is assumed at this point that the participating facility is now ready to extend the operation of the MAP to its regular workload.

Phase IV

Having thus established measurement comparability, the MAP user can then in principle operate independently of NBS. As long as there is no indication of a loss of statistical control of the process, no further checking with NBS should be necessary. Most participants request a recheck of the starting standards every several years to make sure that no undetected long term drift has taken place. Indeed, this periodic checking serves to increase confidence that the measurements are correct. When the participant desires continued NBS involvement, the NBS coordinator will continue the liaison role.

Each set of measurements performed by the participating laboratory during Phase IV is evaluated by that laboratory with respect to the previously-determined process parameters. The control charts must be kept up-to-date, and new estimates of the process parameters must be made periodically. When NBS is involved in this phase, the participant is notified by NBS whenever data suggest that the process is out of control; the measurements made by the participant since control was last demonstrated must be repeated.

If new weighing designs or procedures must be devised to calibrate nonstandard weight sets (e.g., a Troy weight set), the coordinator will do whatever is necessary to assist the participating laboratory in designing suitable data sheets, and establishing additional control charts. For work that differs from the items normally calibrated by the participant, NBS can provide consulting help and assistance such as might be necessary to accommodate a greater range of weights, calibrate pound standards, and extend pound standards to very large weights normally associated with force measurement. Although the usual mass MAP service uses two one-kilogram masses as the starting standards, the program is sufficiently flexible that the same methods can be used with other mass values.

All mass MAP participants do not have the same uncertainty requirements. The MAP determines the actual uncertainty of the participant's measurement process. Assuming the participant's mass laboratory and procedures are of reasonably good quality, random errors tend to predominate. At the time of this writing, mass MAP participants who have chosen to involve NBS in Phase IV achieve total uncertainties ranging from a few parts in 10^7 to a few parts in 10^5 for 1 kg.

For best results, the participant should have high quality balances, barometers, thermometers, and hygrometers, and a reasonably good quality mass laboratory environment, that is, one with minimal air flow, including minimal horizontal thermal gradients and a positive vertical temperature gradient. In the usual program, the participant must have two one kilogram weights plus a set of sensitivity weights ranging from about 1 milligram to 500 milligrams. Calibrations are scaled up or down from this starting set. These and the test set check standards should be good quality nonmagnetic stainless steel weights with a highly polished surface finish. The NBS coordinator can provide advice on selecting suitable standards and equipment.

The fee for new participants typically ranges from $2,000 to $4,000 depending on the technical requirements of the participant, the particular weights utilized in the process, etc. The fee covers calibration of the starting standards, consultation by NBS staff, and all other tasks required to complete Phase III. When the customer requests NBS analysis of the check standard data during Phase IV, fees are $200-$300 per data set depending on the quantity of data submitted.

NBS processes the mass MAP data using a computer program written in Fortran V. This program takes into account the air buoyancy correction for mass calibration. The software has provisions for the use of different weighing methods and different designs for the intercomparison of weights in a given set. The method of least squares is used to determine the values of the weights and their variances. Statistical tests are provided to monitor the precision of the measurement process. The software consists of approximately 3700 lines of FORTRAN code, and requires approximately 20K words of memory. Documentation for this program and a listing are contained in NBS Tech Note 1127 (Bibliography item number 27) by Varner et al. A magnetic tape containing this program is also available from NBS for a fee. Contact Ruth Varner for details: (301) 921-3651 or FTS 921-3651.

5.2 D. C. Voltage (Standard Cells) (SP 250 No. 3.4B)

The purpose of this MAP service is to assure the accuracy of dc voltage measurements at the one volt level. The transport standard for this program consists of a standard cell enclosure containing four cells, equipped with constant temperature control and packaged for shipment by air freight (Figure 4). At present NBS has 12 of these transport standards. Eight are maintained at a constant temperature during transit by means of battery packs connected to the twelve-volt emergency power terminals of the enclosures. The packs also contain battery chargers, and are designed to power the enclosures in the laboratory. In this way the effects of line disturbances on an enclosure are minimized. The remaining four enclosures contain built-in batteries. All 12 enclosures are capable of performance at the 0.2 - 0.3 part-per-million (ppm) level of reproducibility, even after having been frequently transported.

The measurement uncertainty achieved in this MAP is limited primarily by random error. The major components of the random error are: (a) day-to-day fluctuations in temperature-corrected cell emfs caused by temperature hysteresis effects, (b) the finite resolution of the NBS measurement apparatus, (c) thermal emfs, unstable with time, which occur in the measuring circuit due to room temperature changes and drafts, (d) temperature coefficients of the enclosures as a whole, (e) lack of

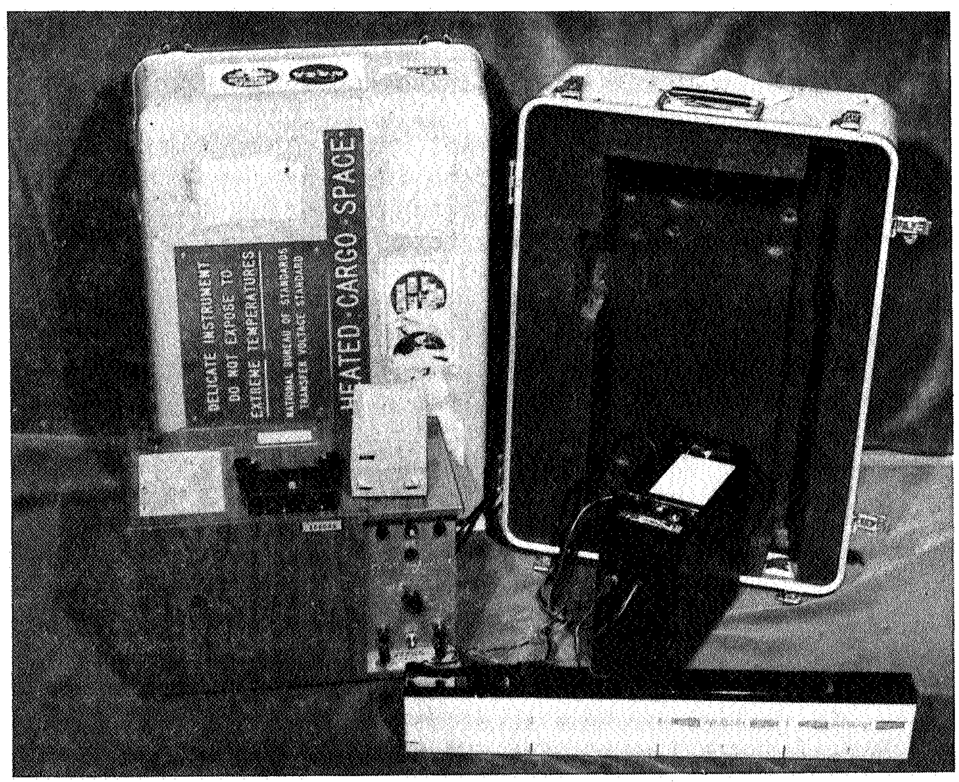

FIGURE 4. STANDARD CELL TRANSPORT PACKAGE.

resolution or instability of the apparatus used to monitor the cell temperatures, (f) changes in temperature gradients or in enclosure temperature due to atmospheric "pumping" of cool air into the enclosure when the barometric pressure changes, or vibration effects on the control circuitry, (g) controller irregularities due to power line transmitted interference, (h) effects of electrostatic or electromagnetic pick-up on the measuring system, (i) detector drift, and (j) the temperature upsets caused by small electric currents passing through the cells.

The experimental designs prescribed for intercomparing the cells average "left-right" or offset errors. One potentially significant source of systematic error not corrected for or quantified in the transfer is that resulting from errors in the divider ratio or scale of the potentiometer used to measure the differences in cell emfs. If, for example, the voltage divider in the potentiometer consists of a 10,000 ohm resistance in series with a 1 ohm resistance, the actual resistance ratio will generally not be exactly 10,000 to 1. If the potentiometer has been carefully calibrated so that the actual ratio has been determined to be, say, 1.00003 or 0.999998 times the nominal ratio, this constant multiplier correction factor can be incorporated into the measurement. If the actual ratio is not known, then the uncertainty in the divider ratio contributes an additional amount to the total uncertainty.

This scaling error can become a particularly significant source of error if the temperature of the customer's cells differs from that of the transport standards. For example, if two otherwise "identical" cells differ in temperature by 2 °C, their voltages will differ by about 80 microvolts. Since the actual divider ratio in the potentiometer may deviate from the nominal value by as much as one percent, a systematic error of the order of one part per million due to an uncorrected divider ratio is possible. This error is propagated when cells are calibrated by the participant at temperatures other than that of the NBS MAP standard.

If the MAP participant has a good quality potentiometer that has been accurately calibrated, the uncertainty of a single transfer with the standard cell package transport standard is typically of the order of 0.5 ppm or better. To realize maximum accuracy for a voltage MAP, scaling error should not exceed 0.1 microvolt. A method for checking scaling error can be provided by NBS to MAP service users.

In this MAP service NBS provides detailed instructions to the participants concerning how the measurements are to be done. The participant must have good quality in-house standard cells and a calibrated potentiometer capable of making intercomparisons at the 0.1 microvolt level. Therefore, when a new participant (or group of participants) expresses a desire to participate in the voltage MAP, NBS requests a complete description of the participant's calibration equipment. (For example, potentiometers employing slidewires for the least-count dial pose operating difficulties due to thermal emfs generated at the wiper during the balancing operation and are therefore not acceptable for use in the MAP.) NBS will provide consulting assistance to the participant as needed to resolve any problems that may arise.

NBS staff advises new participants on how to maintain control charts for their own standard cells in cases where the participants are not already doing so. Beginning October 1, 1984, NBS will require that MAP participants maintain control charts.

Because the uncertainty requirements vary from one participant to the next, no particular frequency of intercomparison is recommended. Many participants find that as they gain experience in the program, the interval between transfers can be increased with no loss in accuracy.

The transport standard is normally kept by the participant for about four weeks; data analysis and issuance of the test report by NBS following the return of the transfer standard to NBS takes four to five weeks.

Copies of NBS computer programs for dc voltage, and to a more limited extent, for resistance and capacitance MAPs are available. Tapes are available for dc voltage MAP data analysis on the HP9830* or HP9845A* (BASIC), but are not well documented. Punch cards with FORTRAN are available for IBM*, Univac*, or Burroughs* machines. A listing of a BASIC program is also available. A magnetic tape and listing of a FORTRAN 77 least-squares routine with zero weighting capability can be obtained from NBS. These programs are for standard cell data analysis. For resistance and capacitance measurements, no general programs are available, however, NBS can provide a general routine on cards for a matrix inversion needed for these programs.

A FORTRAN program called "Analysis" is used for control charts and computation of MAP results. It performs a linear regression on weighted data (including zero weights). A listing of this program and a sample input/output set are available. There is no detailed software documentation currently available.

Contact Norman Belecki for details: (301) 921-2715 or FTS 921-2715.

5.3 Resistance (SP 250 No. 3.1B)

This MAP service is provided to quantify the participating laboratory's uncertainty for resistance measurements at one or more decade values of resistance in the range from 1 ohm to 10^9 ohms. Measurements at two or more selected levels also provide a test of the laboratory's ability to scale or make ratio measurements linking those levels.

The transport standard (Figure 5) consists of at least three standard resistors for each denomination requested, except for the multi-megohm range in which only one resistor is used. Four-terminal resistors are used at resistance levels of 10^4 ohms and below; two-terminal resistors are used above 10^4 ohms.

---

* Mention of specific computers is for identification purposes only and does not imply endorsement by NBS.

Experience at NBS has shown that standard resistors of the type used for transport standards follow a resistance vs time relationship that begins as an exponential curve but decays to a straight line with small but constant slope a few years after manufacture. In the 1 ohm to 10^5 ohm range this slope is typically less than 1 ppm per year. (This initial decay is presumed to arise from room temperature annealing that takes place following manufacture.) Since this behavior is well established, all available historical data on each resistor are used by NBS in establishing its value at any given time. When a standard returns to NBS following a transfer, at least eight new data points are obtained and combined with all previous data to compute a curve that determines the resistance value of the standard during the time it was in the participant's laboratory. Experience suggests that a good-quality standards laboratory should be able to achieve an uncertainty ranging from less than 0.1 ppm at the one ohm level to 20 to 30 ppm at the highest resistance levels, assuming, of course, that a suitable in-house measurement control program is employed.

The level of performance of some typical NBS transport standards is presented below in terms of the residual standard deviation of the fit of several years' data to a straight line. (Part II of this guide explains what is meant by a residual standard deviation.)

Nominal Value of Resistor (Ohms)	Range of Residual Standard Deviations (PPM)		Minimum No. of Measurements per Resistor
	Low ("Best" Resistor)	High ("Worst" Resistor)	
1	0.035	0.124	200
1×10^2	0.035	0.089	45
1×10^3	0.051	0.086	30
1×10^4	0.040	0.182	50
1×10^6	0.53	1.61	22
1×10^8	2.0	3.3	25

The principal component of the total uncertainty of a transfer for resistance values of 10^6 ohms or less is random error, resulting from short-term effects of temperature, pressure, humidity, and the finite resolution of the measurement instrumentation. For values above 10^6 ohms, large systematic components of uncertainty exist in NBS values; it is at that point that a transition from a well-controlled environment of stirred oil at 25 °C to air at a laboratory ambient of 23° C is made. These systematic sources of error include errors in the temperature coefficients of the resistances used, in making the actual temperature measurements, and in leakage or shunting of the resistors by less-than-ideal insulators.

More deviation from the specified procedures can be tolerated in the resistance MAP than in the direct voltage MAP. There are many acceptable techniques for making resistance intercomparisons and maintaining resistance standards. By mutual agreement, specific intercomparison techniques may be tailored to the equipment, operating procedures and need of each individual laboratory. Only general instructions and data reporting formats are provided by NBS. As in most other MAPs, NBS does

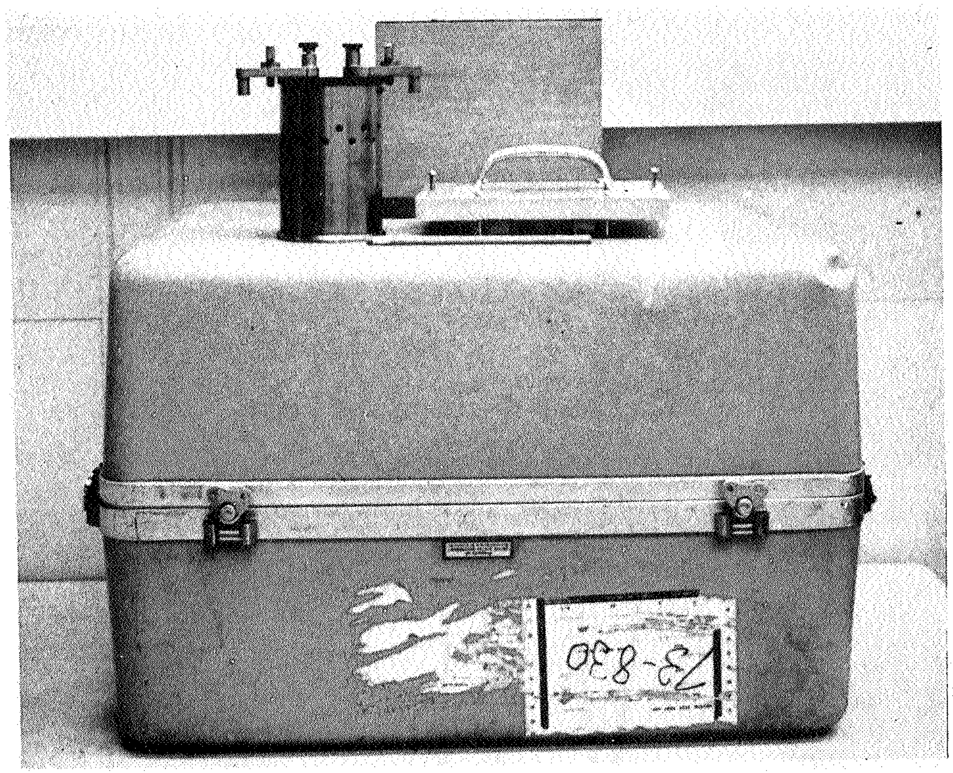

FIGURE 5. RESISTANCE TRANSPORT STANDARD AND SHIPPING CONTAINER.

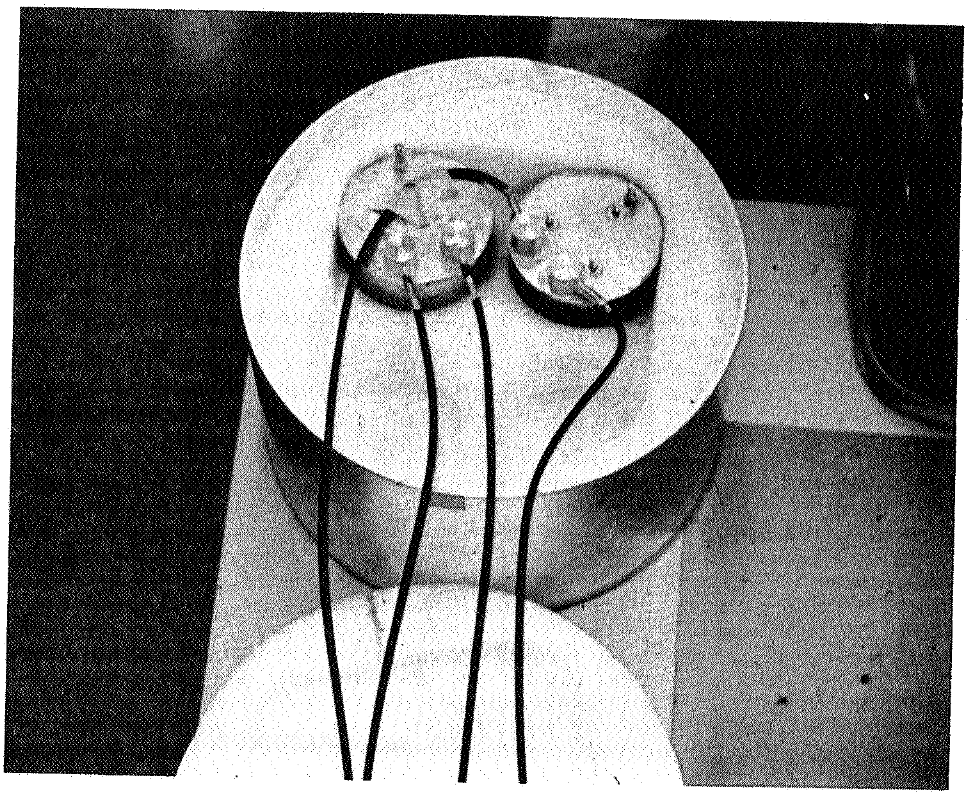

FIGURE 6. CAPACITANCE TRANSPORT STANDARD.

not recommend any specific frequency of intercomparison since accuracy requirements vary so much from participant to participant. Typically, the resistance transport standard remains in the customer's laboratory for four to six weeks. Five to six weeks are required for rechecking the transport standards at NBS and preparing the test report.

(See Section 5.2 for information concerning the availability of computer programs for analyzing resistance MAP data.)

5.4 Capacitance (SP 250 No. 3.3B)

The purpose of this MAP is to assure the accuracy of capacitance measurements made by participating laboratories at the 1000 picofarad (pF) level, 1000 Hertz.

The transport standard for this MAP consists of four nitrogen gas dielectric standard capacitors, each nominally 1000 pF, which have been modified by NBS (e.g. trimmer capacitor removed, special low leakage head installed) to permit the highest possible precision to be achieved (Figure 6).

These transport standard capacitors exhibit small but significant shifts in value when subjected to mechanical shock. They also display small hysteresis effects when cycled in temperature. These effects are negligible during normal usage. When these capacitors are used for the MAP transport standard, however, they must be protected from shock and continuously maintained at a constant temperature. The shipping container is equipped with a battery pack to maintain a constant temperature (approximately 30 °C) during transit, and an ac power supply to maintain the temperature while the standard is in the participant's laboratory or at NBS.

Typical uncertainties for capacitance MAP transfers are of the order of 0.7 ppm, although some laboratories do better. The user, of course, must have an ongoing measurement control program in order to be able to determine the total measurement uncertainty. The uncertainty consists of random errors due to small fluctuations in temperature and both systematic and random errors attributable to the bridge used to intercompare the capacitors. Unless the participant utilizes measurement practices designed to eliminate ground loops, stray capacitances, etc., systematic errors of up to 100 ppm can affect the measurements. When these sources of error are minimized, the uncertainty of the transfer is generally better than 1 ppm. If the participating laboratory uses a high-quality capacitance bridge and the recommended measurement techniques, it should be possible to scale to other capacitance values and only degrade the accuracy by only an additional 1 ppm or so.

The transport standard capacitors are directly compared at NBS with specially constructed standard capacitors. These capacitors are in turn compared to the NBS calculable capacitor (which is the basis for realizing the Farad). The intercomparisons are made with a high-precision transformer ratio-arm bridge developed by NBS.

During the two to three weeks in which the transport standard is in the participant's laboratory, at least eight measurements are made on the capacitors. A minimum of eight measurements are made on the transport standard before it leaves NBS and after it returns to NBS. Follow-up measurements and the preparation of the test report by NBS take approximately two to three weeks. The experimental design and the method of data analysis are essentially the same as for the dc voltage MAP. However, measurements made at NBS involve scaling in 10 to 1 ratios. Thus, the computer software used for this analysis is more general than that used in the dc voltage case. (See Section 5.2 for information on the availability of computer programs for this MAP.)

Standard capacitors do not behave as predictably as transport standard resistors or standard cells. The capacitors do not exhibit a long term drift up or down; instead, they tend to display sudden unpredictable shifts in capacitance of the order of a few tenths of a ppm. As mentioned earlier, these shifts can be minimized by controlling their temperature and by avoiding mechanical shocks. Because of this erratic behavior, however, no long-term curve fitting is done; the value of capacitance during the time the standard is in the participant's laboratory is taken to be the arithmetic mean of the capacitance values obtained before it left NBS and after it was returned.

5.5 Electrical Energy (Watthour Meters) MAP (SP 250 No. 3.7D)

The purposes of this MAP service are to provide a means for transferring the unit of electrical energy from NBS to laboratories that calibrate watthour meters and to provide a means for determining the participant's measurement uncertainty. By participating in this program, meter manufacturers, electric utility companies, public utility commissions, universities, etc., can demonstrate that their electrical energy measurements are compatible with national standards.

The transfer standard provided by NBS consists of a well-characterized commercial watthour meter in a special shipping container. Since the measuring instrumentation in the participant's laboratory may vary, three different meter types are available as transfer standards. The participant specifies which of these three transport standards is appropriate.

Only meters designed for 60 Hz, 120 or 240 volt operation are included in this MAP service. Measurements are made at five amperes, at unity and 0.5 power factor (current lagging or leading voltage). The transfer standard is calibrated at NBS before and after the transfer to the participant. The participant makes eight measurements on the transport standard over a period of about one week. Final measurements at NBS and preparation of the test report can usually be completed in one to two weeks.

Sources of possible error include:

Temperature - Even meters of the same type do not necessarily display the same temperature coefficient: In addition, the temperature coefficient may vary with the power factor. (See references 33 and 34 in the Bibliography for typical data.) All measurements should, therefore, be performed at a temperature of 25 °C ± 2 °C. A thermometer for the participant to use is provided with each transfer standard.

Voltage - The test voltage should be set using a stable voltmeter having an uncertainty of less than a few tenths of a percent. The test voltage should be a sine wave free from appreciable harmonic content.

Magnetic fields - Systematic errors can be introduced by the presence of stray magnetic fields. For this reason, the meter should be tested away from transformers, structural steel girders or other magnetic materials, and should be physically separated from any electrical devices that may generate magnetic fields.

Resistance in the voltage circuit - Precautions should be taken so that no appreciable resistance is introduced into the voltage circuit.

Switching transients - Meter current should be increased and decreased slowly (e.g., with a variable transformer) to avoid saturating the core.

Power factor error - The phase angle should be measured as accurately as possible so that the power factor uncertainty can be minimized.

Measurement time - Measurement time should be at least 100 seconds to achieve the necessary resolution and minimize uncertainties in measurement time resulting from starting and stopping the measurement.

The overall accuracy achievable depends, of course, on the precision of the participant's measurement process and the care taken to eliminate systematic errors. The process precision must be monitored over a period of time using in-house check standards. The typical random error (three standard deviations) for the NBS transport standard is about 0.01 percent; the overall uncertainty of the transport standard value is about 0.02 percent. Most participants having a well-equipped standards laboratory should be able to achieve an uncertainty of the order of 0.05 percent or less in their own measurements.

A BASIC program for analyzing watthour meter MAP data called "ENMAPR" consisting of approximately 400 lines of code is available for a 16 bit Interdata* computer. This program analyzes both NBS data and the participant's data. Registration and test conditions (voltage, current frequency, temperature, etc.) are input; and means, corrected means, and standard deviations are calculated for both NBS measurements and those of the participant. A test for statistical control is made using Student's T distribution, and the offset of the participant's calibration process relative to NBS is calculated, along with the total uncertainty. Copies of the program listing are available, and copies of the program on floppy disks can be made available for those with Interdata computers. It is anticipated that the software will be rewritten for a new computer system within the next year. Contact John Ramboz for details: (301) 921-3121 or FTS 921-3121.

* Mention of specific computers is for identification purposes only and does not constitute endorsement by NBS.

5.6 Temperature (SP 250 No. 7.3J)

The purpose of this MAP service is to assure the accuracy of the calibration of temperature standards made by participating laboratories in the -183 to +630 °C temperature range when using platinum resistance thermometry. Special arrangements may be made if participants are interested in only a portion of this temperature range.

The MAP transport standard consists of a set of three commercial glass-sheathed standards-type platinum resistance thermometers (SPRTs) packaged in a special shock-proof shipping container (Figure 7). (The SPRTs are delicate instruments; mechanical shock or certain heat treatments may result in shifts in calibration.) Three SPRTs are used to assess both the reproducibility and the accuracy of calibrations performed by the participating laboratory. Typically the participant is not informed of the resistance vs temperature relationship of these SPRTs ahead of time, hence, the SPRTs are "blind samples."

MAP participants should use the techniques in Monograph 126 (Bibliography item 35) and the same fixed points as in the NBS calibrations, or an SPRT previously calibrated by NBS. In order to achieve high accuracy, SPRTs used as standards should be of the matte-finish type to avoid systematic errors arising from light pipe effects in the glass sheath. The lab must have a triple point cell and a calibrated resistance bridge.

Instructions are provided to the participant for packing and unpacking the SPRTs. Each thermometer is measured by NBS and inspected prior to shipment. After unpacking the transport standard (and visually inspecting it for damage) the participant measures the resistance of the SPRT at the triple point of water using the triple point cell. A preliminary check of the resistance at the water triple point is used as a "go-no go" check to ensure that the thermometer has not been damaged in shipment.

This measurement is reported by telephone to NBS; if this value is consistent with the data taken by NBS before shipment, the participant proceeds with further measurements. The measuring current is normally 1 mA, but tests are also made at 2 mA or $\sqrt{2}$ mA to quantify heating effects.

After annealing the SPRT, calibration of the transport standard is carried out. Data are taken by NBS and the participants at the fixed points defined in the International Practical Temperature Scale (IPTS-68) using the procedures outlined in NBS Monograph 126. NBS provides a worksheet on which the participant can record the data. The participant then calculates the thermometer constants from the experimental data, records them, and prepares a table of resistance vs. temperature.

The SPRTs are recalibrated upon return to NBS and the data are compared to NBS's calibrations. NBS provides a plot of the participating laboratory's temperature deviation from NBS values and a written analysis of the data including any pertinent observations. The participants are encouraged to perform the measurements without taking any special precautions above and beyond those normally used in the laboratory's routine calibration activities. In this manner, a realistic estimate of the

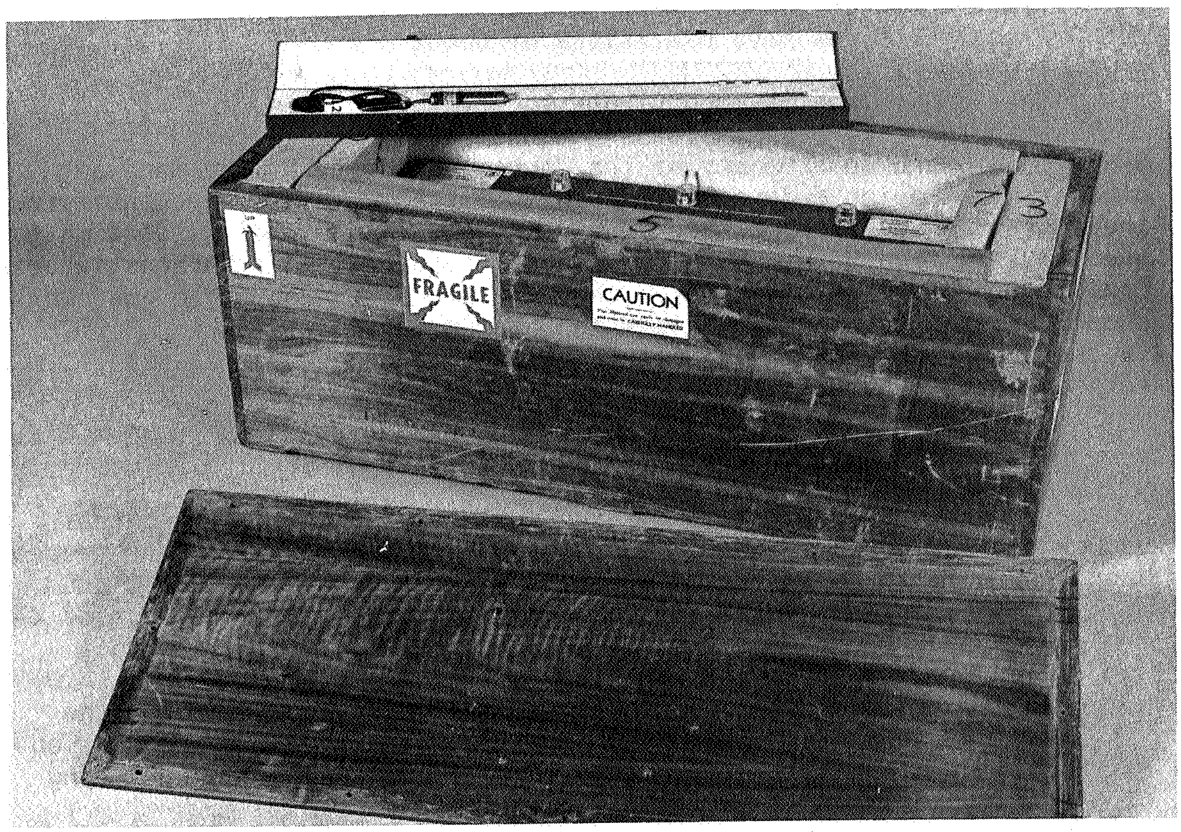

FIGURE 7. PLATINUM RESISTANCE THERMOMETER TRANSPORT PACKAGE.

laboratory's measurement uncertainty is obtained. In a typical transfer, the participant makes several measurements over a period of two to three months. Typical turn-around times are three to four weeks from the time NBS receives the participant's data until a test report is sent to the participant.

Measurements on the SPRTs are made at NBS using a dc current comparator bridge or other high-precision four-terminal measurement instruments. Adjustments are made to the measured values to compensate for the deviation of the actual measurement conditions from the defined conditions. These include the hydrostatic head correction for the height of the fixed-point liquid, the external (atmospheric) pressure of the fixed-point liquid, etc. NBS quantifies the effect of these factors through experiments in which these factors are intentionally varied. The best NBS SPRT calibrations have precisions of about 0.1 to 0.2 mK.

Sources of error that may contribute to the total uncertainty include changes in the calibration of the measurement instruments, changes in the SPRT itself, and uncertainty of the degree of purity of the materials used as fixed-point references (e.g. zinc). As a result of quantifying these sources of error, NBS currently assigns an uncertainty of 1 mK to the values assigned to the MAP transport standards. A standards laboratory conscientiously participating in this MAP and having suitable equipment should be able to come very close to this uncertainty figure. Participants in the temperature MAP have uncertainties that range from close to 1 mK to hundredths of a kelvin.

No rigid recommendations can be given concerning how often a lab should utilize the temperature MAP service. Experience has indicated that a lab whose temperature measurements are in a state of statistical control using in-house check standards and control charts to monitor the process should be able to go at least three years between transfers from NBS without significantly degrading the confidence in the correctness of the measurements.

Based on comparisons with other national standards laboratories, the agreement of temperature measurements among the national standards labs of the principal industrialized countries is comparable to that among high quality standards labs within the U.S.

Precision Thermometry Seminars are conducted at NBS twice each year (usually in March and September) to help standards laboratory personnel become familiar with good laboratory practices for achieving accurate temperature measurements with SPRTs, thermocouples, and liquid-in-glass thermometers. These seminars include a discussion of the measurement procedures for SPRTs upon which the MAP services are based.

NBS is also investigating the feasibility of temperature MAPs at lower levels of accuracy using metal-sheathed PRTs of the type used in industry. As part of this study, a round-robin measurement program is in progress under the auspices of the American Society for Testing and Materials (ASTM), Committee E20.03 on Resistance Thermometry.

A short FORTRAN program is available for the calibration of platinum resistance thermometers (PRTs). This program inputs the constants provided by the customer, calculates the resistance vs. temperature for the PRT, and compares these data to NBS's data for the thermometer. Copies of the program listing can be made available, and copies of the magnetic tape containing the program can be provided for a fee. Contact George Furukawa for details: (301) 921-2742 or FTS 921-2742.

5.7 Laser Power and Energy (SP 250 No. 4.4B)

The units of laser power and energy are realized in the U.S. by means of three types of isoperibol (constant temperature environment) calorimeters that compare absorbed laser radiation to an equivalent quantity of electrical energy. These calorimeters, constructed and maintained by NBS are part of a laser measurement system used to calibrate other laser power and energy meters to an uncertainty of about one to five percent depending on the power (or energy) and the wavelength at which the calibration is performed. The MAP for laser power and energy uses one or more well-characterized power meters or calorimeters as transport standards to assess the participant's measurement process for laser power and energy. In addition, the participant makes regular measurements on one or more in-house reference power meters or calorimeters (check standards) in order to establish the random error and to provide assurances that the process is stable.

The services currently available are given in the table below.

Laser Power and Energy MAP Services as of 1984

Wavelength of Laser	Power or Energy Level
514.5 nanometers	10-600 milliwatts
632.8 nanometers	1 milliwatt
632.8 nanometers	1, 30, or 100 microwatts
647.1 nanometers	10-200 milliwatts
1.06 micrometers	10 milliwatts-1 watt
1.06 micrometers	Q-switched, 100 millijoules-10 Joules
10.6 micrometers	5-50 watts

Each of the three laser calibrating systems used to support the NBS laser MAP service employs the two calorimeter-beamsplitter configuration.* All the significant parameters for each calorimeter, including the electrical calibration coefficient and the absorption and window transmission for each laser wavelength being used, have been evaluated to relate laser energy to electrical energy. The application of theory to the voltage data (proportional to temperature) taken at equal time intervals is accomplished by using a special least squares computer program.

A properly designed and operated calorimeter can make a valid comparison of energies independent of the time (up to five minutes) required to put the energy into the calorimeter. For instance, it is valid to compare the energy in a pulse to continuous wave (cw) energy, applied over a five minute period, or to compare a laser pulse to an electrical input of 10 to 300 second duration.

Using a beamsplitter configuration, energy measuring calorimeters can be calibrated independently of laser stability, either with cw power or the energy of single or multiple pulses. In addition, if accurate timing and a stable laser are employed, laser power meters can also be calibrated.

Diagrams showing the construction details of the calorimeters can be found in the articles in the Bibliography (e.g., items 39-41). Laser sources at NBS used with these calorimeters are carbon dioxide, argon, krypton, helium-neon and neodymium-doped YAG cw lasers, and a pulsed YAG laser with a pulse energy of about 0.1 joule and a pulse width of about 30 nanoseconds.

Since laser power and energy are highly-derived quantities, more sources of error are possible than typically encountered in MAPs for measurements closer to the base SI units. Sources of error which must be considered in laser power and energy measurements include:

Beam size - The beam size should be small enough that the various instruments can capture the total beam, and large enough that the energy density of the beam does not damage the surface of the absorber.

Beam alignment - Proper alignment ensures total beam capture and avoids interference problems with reflections.

Window transmittance and beam splitter ratio - These wavelength-dependent ratio measurements must be determined for each laser wavelength of interest.

Electrical measurements - To relate the laser energy to base SI electrical units, time intervals, voltages and currents of the heater circuitry must be measured accurately with respect to national standards.

Long term drifts in the parameters characterizing the measurement system - Examples of parameters that may change with time are:

---

* The equipment and techniques discussed in this section are explained in more detail in Bibliography items 38-45.

- electrical calibration coefficients
- absorbing surface of the calorimeters
- window transmittance

Control charts for the beamsplitter measurements with the two calorimeter-beamsplitter configuration are used at NBS to verify the stability of the calibration systems.

Generally the transfer standards used in the NBS laser power and energy MAP have their own digital readout and do not require special equipment or a computer program to reduce the data. If the participant uses isoperibol calorimeters in a calibration configuration similar to the NBS system, computer programs are available from NBS for reducing such isoperibol data. Whether or not the specialized NBS data analysis computer programs can be utilized in other cases by a MAP participant depends on the particular calibration configuration. Listings are available for several BASIC programs related to laser power and energy measurements. Most run either on a large CYBER* computer or on a Tektronix* 4052. For example, one program calculates beam splitter attenuation. The refractive index for the wavelength of interest, the wedge angle of the beam splitter, etc. are input, and the attenuation of the transmitted and reflected beams are calculated. A second program inputs calibration factor and other calorimeter data and outputs the energy. This program is sufficiently general to apply to most laser calorimeters. The basis for these programs is described in the references listed in the Bibliography, particularly items 38 and 44. Contact Aaron Sanders for details on the available programs and their applicability: (303) 477-5341 or FTS 320-5341.

5.8 Gage Blocks

The MAP service for gage blocks is intended for laboratories that need to document, on a continuing basis, the measurement uncertainty of their gage block calibration process. Check standards, redundant measurements, and statistical analyses form the basis for monitoring the in-house calibration process. Offset from the defined unit of length is determined from the participant's measurements on transfer standards from NBS. Two eighty-eight piece sets of steel gage blocks, spanning the range from 0.100 inch to 4.0 inches, are available as transfer standards. (Check with NBS regarding the availability of metric transfer standard sets.) Detailed instructions for operating the measurement assurance program under three different options are given in Monograph 163 (Bibliography item 46). The three options allow for varying degrees of rigor in operating a MAP, and the choice of option depends on a number of factors including: 1) availability of master standards; 2) availability of time to operate the program; 3) equipment; and 4) accuracy requirements. In principle, the participant can choose to operate the MAP using any of the three options, but only the second option is fully supported by NBS. Briefly, the three options are as follows:

*Mention of specific computers is for identification purposes only and does not constitute endorsement by NBS.

Option 1 is best suited for interferometric measurement processes. One master set of steel blocks and 10 steel check standard blocks are provided by the participant. The process is monitored through the incusion of check standards in the calibration workload at roughly equal size intervals. Transfer standards of the same nominal size as the check standards are provided by NBS. The analysis of data is the responsibility of the participant.

Option 2 is intended for electro-mechanical comparators or interferometers used as comparators. Two master sets of steel gage blocks are required. Duplicate measurements are achieved by comparing each test block with each master block, and process control is maintained on the difference in observed length between master blocks of the same nominal size. Two sets of transfer standards are furnished by NBS, and complete data analysis along with a report of test is furnished by NBS.

The program for Option 2 progresses in four stages. In the first stage, a database created from the participant's calibration workload is used to establish parameters for the quality control aspect of the program that is to follow. This database requires a minimum of six experiments. In each experiment a complete set of gage blocks is compared to the two master gage block sets, and process parameters are computed from the resulting data.

In the second stage, the transfer standards from NBS are measured. This experiment is performed twice on each set of transfer standards. Values and uncertainties are assigned to the master sets only if these experiments are in statistical control as judged by the process parameters. Therefore, the participant is urged to take care that the first stage of the program is truly representative of the measurement process so that problems with the transfer can be avoided.

In the third stage, the data from the transfer experiments is incorporated into the database, and the process parameters are updated. In the fourth stage, the participant proceeds with the calibration program using MAP procedures and the established process parameters until another transfer with NBS is scheduled.

Option 3 is intended for electro-mechanical comparators. Two master sets of steel gage blocks are required. Option 3 is a more complicated procedure than Option 2 as it involves comparisons among two test blocks and two master blocks of the same nominal size. Process control is maintained on the difference in length between two master blocks of the same nominal size as calculated from the design. Control is also maintained on process variability as computed from each set of intercomparisons. Two sets of transfer standards are furnished by NBS. Data analysis is the responsibility of the participant.

The purpose of the MAP is to maintain a continuous check on the state of the calibration process thereby guaranteeing that measurements are properly related to the defined unit of length within the stated uncertainty. This implies that the daily calibration procedure is identical to the MAP procedure. Because considerable time and effort may be required to implement a transition from single measurements to a MAP procedure and to establish the database for the program, the perspective participant is urged to weigh potential benefits of the program against the investment of time and effort before committing to the program.

Charges for the MAP service are on an at-cost basis since the degree of NBS involvement varies depending on the option and number of gage blocks selected.

Computer software is available for gage block calibrations in which an unknown set of gage blocks is calibrated against two standard sets, with control being achieved by monitoring the difference between the two standards. The software has provisions for establishing process parameters, testing for process control, assigning values to test blocks with associated uncertainties, and updating process parameters. The published version of the software consists of approximately 4000 lines of FORTRAN code and requires approximately 26K words of memory. Documentation for this software is contained in NBS Tech Note 1168 (Bibliography item number 53) by R. Varner. A magnetic tape containing this program is available from NBS for a fee. Contact Ruth Varner for details: (301) 921-3651 or FTS 921-3651.

6. Regional or Group MAPs

MAP services are staff-intensive for NBS when each participant is served separately. In order to serve more users of MAP services in the future without a proportionate increase in NBS staff, a more efficient way to disseminate MAP services has had to be found. The regional or group MAP addresses this concern. The Measurement Assurance Committee of the National Conference of Standards Laboratories (NCSL) has been active in encouraging laboratories to form regional groups and establish measurement assurance programs. (See Chapter 8).

In a group MAP, a number of participating laboratories, sufficiently close together to permit regular personal contacts; e.g., Southern California, New York/New Jersey metropolitan area, etc., work together to achieve traceability to NBS. NBS interacts with the group as a whole through one individual designated as group leader and in some cases, one laboratory designated as the "pivot laboratory."

In a group MAP each participant's standards are compared directly with the NBS transport standard. Each participant receives a test report from NBS and therefore, each group MAP participant achieves primary or direct traceability to NBS.

Normally the MAP group leader is responsible for scheduling the circulation of the NBS or other member's transport standards among the participants, assembling the data, and in some cases analyzing the data. The pivot laboratory's responsibility may be a permanent designation, or may be a rotating assignment, with each member in turn acting as pivot laboratory. If one of the laboratories in a group is clearly superior in facilities and personnel, it may be appropriate for this laboratory to be the pivot laboratory on a permanent basis (if agreeable to all group members). The role of the pivot laboratory may then involve providing more consultation and leadership to the other group members than in the case where all participating laboratories in the group including the pivot laboratory regard each other as equals. The group leader for the MAP is generally a person from the pivot laboratory, but this is not a prerequisite. (Some group MAPs may not use a pivot laboratory, but would still have a group leader.)

In order to efficiently serve a large number of laboratories in many measurement areas, the amount of data analysis by NBS has had to be minimized. NBS makes computer programs for data analysis available to the group. These programs range from those requiring large computers to those for desktop calculators. The delegation of much of the operation of the program to the participants does not mean that accuracy suffers, since the MAP provides the same accuracy whether NBS analyzes the data or the participants do.

The regional approach to measurement assurance is now being used with great success in several parts of the U.S. for disseminating the volt. Additional laboratories in other parts of the U.S. are now forming voltage MAP groups and interest is increasing in extending this approach to other types of measurements. Further experience may indicate that the group approach does not lend itself well to some types of measurements, and for such measurement quantities the "one-on-one" approach will continue to be used. The group approach may not be feasible when laboratories are engaged in highly competitive proprietary activities. The trend will, however, undoubtedly be towards increased use of regional MAPs wherever feasible.

As an example of how a group MAP operates, the operation on a regional basis of the dc voltage MAP (The Volt Transfer Program), is described in the remainder of this section. The Volt Transfer Program (the MAP for standard cells) is designed to quantify the uncertainty of the assignment of the unit of voltage at each participating laboratory in the group.

As the first step in the establishment of a Regional Volt Transfer Program group, the NBS coordinator will review the procedures and equipment used by each participant. Based on the results of this analysis, the coordinator may urge certain participants to carry out internal experiments to verify the suitability of their equipment and procedures. The coordinator may also recommend changes in some operating procedures and may suggest procedures for monitoring the local measurement process to determine whether or not it is in a state of statistical control. When measurement problems are encountered, the NBS coordinator will assist in solving these problems. As can be seen, this first step in a group MAP does not differ from individual laboratory MAP services.

In the regional Volt Transfer Program, the five or six participants in the group may take turns being the pivot laboratory. In this particular group MAP, the optimal group size appears to be five or six labs; in other group MAPs this number may differ.

When NBS and the group have agreed on a schedule for the intercomparisons, the NBS transport standard is sent via air freight under carefully controlled conditions to the pivot laboratory. The pivot laboratory compares this transfer standard to its reference group of cells using the procedures for intercomparison recommended by NBS. (The pivot laboratory calibrates the transfer standard.) Each participating laboratory checks its transport standard against in-house reference standards before taking it to the pivot laboratory and again after it is brought back from the pivot laboratory. The group members bring their transfer voltage standards to the pivot laboratory so that intercomparisons can be made with the NBS standard, with the pivot laboratory's standards and with each others' standards as called for by the experimental design. In the Southern California Volt Transfer Group, all comparisons at the pivot laboratory are made by pivot laboratory personnel (by mutual agreement of the participants). After the measurements are completed, the NBS transfer standard is returned to the Bureau for remeasurement.

When the data from individual laboratory control charts and from calibration comparisons with the NBS transport standard are analyzed by NBS, it is possible to establish the measurement uncertainty of each participant's measurement process. As additional transfers and intercomparisons take place and additional data are generated, the uncertainties have tended to decrease from the original estimates made following the first transfer. Participants in the Southern California regional Volt Transfer Program have achieved an uncertainty considerably better than one part per million.

In the case of standard cells, where the transportability of the standard is a limitation, all group participants take their transport standards to the pivot laboratory where the intercomparisons with the NBS transport standard are made. For other MAPs where the NBS transport standard is rugged and stable (e.g. gage blocks), it may be more efficient for the pivot laboratory to circulate the NBS standard among the group members with each participant measuring the standard in-house. Suitable arrangements for using the standard must be worked out for each group MAP on a case-by-case basis.

The frequency of intercomparison depends on the group's requirements, and typically varies from four intercomparisons per year to one every two years for the regional Volt Transfer Program. Each participant continues to follow the recommended measurement procedures between intercomparisons, keeping control charts and monitoring voltage measurements to make sure they remain in statistical control. Thus, as with MAP services to individual laboratories, several sources of uncertainty inherent in the regular procedure in which customers send or bring their cells to NBS are eliminated. The voltage MAP "calibrates" the entire process (including the technicians, the laboratory environment, etc.), not just a portion of it (the reference cells).

Experience with the original Southern California Volt Transfer Group continues to be favorable. To determine the participant's volt in relation to the group mean, an annual comparison within the group has been made. It has been demonstrated* that the intervals between NBS transfers can be extended from a one-year interval to an interval of two to five years, with an uncertainty of the group mean of about 0.6 ppm. (This is a lower uncertainty than had been observed at the outset.) If experience with this group MAP can be extrapolated to others, it would appear that the cost to the participants in a group can be reduced over time while simultaneously reducing the measurement uncertainty.

Some laboratories may prefer to continue to use the standard cell calibration service of NBS since this is less expensive than the one-on-one NBS MAP service. With the cost of the NBS transfer standard intercomparison in a regional MAP split among the participants, the actual cost to the regional voltage MAP participants doesn't differ appreciably from that of a normal cell calibration by NBS.

---

* See Bibliography item 8.

7. The Future of MAPs

The development of a new MAP service for state-of-the-art measurements is expensive, typically costing several hundreds of thousands of dollars. Once the service is developed, fees paid by the users recover operating costs only. Although the development of a MAP service by NBS at the highest levels of accuracy may be quite expensive, the application of the MAP concept to meet lower accuracy requirements is much less costly and may be quite cost effective. (See Section 8.5.)

In addition to new MAP services developed with Congressionally appropriated funds, NBS has developed new MAP services with funding provided by other agencies. For example, the Laser Power and Energy MAP was originally developed for the Department of Defense, but is now available to the general public.

A sound justification is necessary before development of any new MAP service can proceed. This justification must include evidence that the service is needed and that once developed, will be used. MAP services will not be developed to meet needs that can better be met by the private sector.

It has been asked if NBS intends to replace all calibration services by MAPs. The answer is no. At the present time NBS offers approximately 400 different calibration services. While the number of MAP services available from NBS can be expected to increase during the coming decade, it is extremely unlikely that NBS could ever replace all its calibration services with MAPs because of the high cost involved in developing MAP services. There will continue to be both MAP and calibration services offered for several measurement quantities.

Future trends in MAP services include:

- Increased use of the regional or group MAP wherever feasible;
- Structuring MAP services into a "do-it-yourself" mode without sacrificing accuracy or convenience.
- More computerization of data gathering and data analysis (e.g., transmitting data from participants into an NBS computer in real time over phone lines in order to speed the preparation of NBS test reports for MAPs);
- Expanded documentation of MAP techniques in the form of NBS handbooks and technical notes.

8. The Role of Other Organizations in MAPs

The development of the MAP approach to measurements has involved many people outside NBS. Many organizations have successfully used MAP techniques to solve "real world" measurement problems, have assisted NBS in disseminating MAP services, and have assisted in the development of the MAP philosophy.

8.1 The National Conference of Standards Laboratories

The National Conference of Standards Laboratories (NCSL) is a nonprofit, laboratory-oriented organization that promotes cooperative efforts toward solving the common problems faced by standards laboratories in their organization and operation. It was established in 1961 under the sponsorship of the National Bureau of Standards. Its membership consists of academic, scientific, industrial, commercial, and governmental laboratories involved in the measurement of physical quantities, the calibration of standards and instruments, and the development of standards of practice. It provides liaison with technical societies, trade associations, and educational institutions interested in these activities.

NCSL's Measurement Assurance Committee provides laboratory management with methods for evaluating and improving the quality of measurements performed by their laboratories. This committee has played a key role in increasing awareness within the calibration and standards community of the importance of measurement assurance programs. The committee leadership has devoted considerable time to informing NCSL members of the advantages of the MAP approach from the perspective of the standards or calibration lab manager, and has served as a catalyst in encouraging laboratories to form into groups for regional MAPs. In fact, the concept of the regional approach to MAPs was conceived by a former NCSL President.

Measurement assurance has been a featured topic at numerous NCSL regional meetings, so that there is increased awareness of MAPs and NBS MAP services among the NCSL membership. As a result, interest in MAPs has increased substantially during the past few years. Organizing MAP groups among laboratories within particular regions is a time-consuming activity. In some cases NCSL has taken responsibility for organizing MAP groups so that NBS staff have been able to devote their attention to the technical issues involved, thereby speeding up the process.

The current chairman of the NCSL Measurement Assurance Committee, (name and telephone number available from the NCSL Secretariat) can provide more information on NCSL and its role in MAPS.

NCSL Secretariat
c/o National Bureau of Standards
Boulder, CO 80303
(303) 497-3237

8.2 American Society for Quality Control

The American Society for Quality Control (ASQC) has been providing leadership in the quality control field for many years. Its publications on various aspects of QC and its training courses, particularly in the area of statistical methods for QC can be of great assistance to those who wish to become more knowledgeable about statistical methods used in measurement assurance.

ASQC holds the secretariat for the American National Standards Institute (ANSI) Committee Z1 on Quality Assurance. The purpose of this committee is to develop generic quality assurance standards. This committee's publications in the areas of definitions of terminology, statistical methods, etc., are of considerable importance to metrology in general and to MAPs in particular. In 1979 ASQC and ANSI established within the Z1 Committee a Writing Group for Quality Standards for Calibration Systems and Measurements. This group is developing standards for the measurement community regarding quality assurance provisions for measurements, including MAP techniques similar to those described in Part II of this guide.

8.3 American Society for Testing and Materials

The American Society for Testing and Materials (ASTM) has produced a large number of standards and test methods involving accurate measurements. ASTM Committee E11 on statistics has made many important contributions to the field. An ASTM Committee on Quality Provisions in ASTM Standards (E46) discusses the MAP approach in its guidelines on calibration and measurement accuracy.

8.4 Calibration Coordination Group of the U.S. Department of Defense

The Calibration Coordination Group (CCG) of the U.S. Department of Defense (DOD) was established by the Joint Logistics Commanders to provide effective coordination among the calibration and standards activities of the Army, Navy, and Air Force, particularly in the area of interactions with NBS. Each year CCG technical subgroups review their future measurement requirements with NBS. Where current measurement technology appears inadequate to meet future needs in the calibration and standards area, CCG coordinates the provision of DOD funding to NBS in key metrology research and development areas. Through this mechanism, funding has been provided to NBS for the development of a number of MAP services of interest to DOD. Laser Power and Energy, Voltage, and Resistance MAP services were all developed at least in part with DOD funding arranged through CCG. Once developed, the services have been available to the general public as well as to DOD. It is anticipated that this fruitful cooperation between DOD and NBS will continue in the future.

8.5 The Use of Measurement Quality Assurance Techniques in Industry

Most of the attention given to MAPs has focused on the application of this concept for assuring accuracy at the highest levels of standards. Potentially, the application of simplified but similar techniques at lower levels is of even greater significance. Government and industry organizations have applied these techniques successfully in the "real world"; unfortunately, there are few accounts of these applications in the archival literature (see items 18-20 of the Bibliography).

As described in Bibliography item 18, a large aerospace corporation had stringent requirements for the manufacturing of parts for advanced weapons systems. In most companies instruments used by machinists (micrometers, dial calipers, depth gauges, etc.) are recalled at regular intervals for recalibration by the calibration laboratory. This firm, however, concluded that the traditional approach would not provide the degree of quality assurance that was needed to ensure meeting their program objectives, nor would it be welcomed by the employees, many of whom used personally-owned hand tools. Instead, this firm developed what was referred to as a "Personal Hand Tool Verification Program."

The calibration laboratory constructed several items called "master caliper blocks" (Figure 8) commonly called "pretzels" by those who used them. These were stainless steel artifacts consisting of various grooves, slots, holes, protrusions, etc., that could be measured by a machinist using hand gauging tools. The principal dimensions of these blocks could be adjusted from time to time by as much as 100 microinches so that the machinist would not be able to memorize the "correct" dimensions of any particular block. Each machinist periodically made specified measurements on a block and reported the measurements to the metrology laboratory. The metrology lab measured the blocks before and after they were circulated to the machinists, taking enough data to characterize the blocks.

With this new approach, not only was the adequacy of each gauging tool verified, but machinists who were using their gauging tools improperly were identified. A control chart could be constructed, if desired, for each set of components of the measurement system, (consisting of person, tool, block, environment) to monitor their long term stability. The firm that developed this program concluded that it was simple, economical, and effective. The employees, when provided with feedback on their measurement performance, reacted by taking a much greater interest in maintaining their gauging tools in good condition. Experience indicated that the employees took steps to repair or replace defective tools, and some took home study courses on measuring practice made available to them by the company. The result was steadily improving measurements as the program continued.

It is obviously important to present a program like this to the affected employees in a constructive manner. The measurement checks must not develop into an adversarial relationship in which the employee feels that the company will punish him/her for measuring the artifact incorrectly.

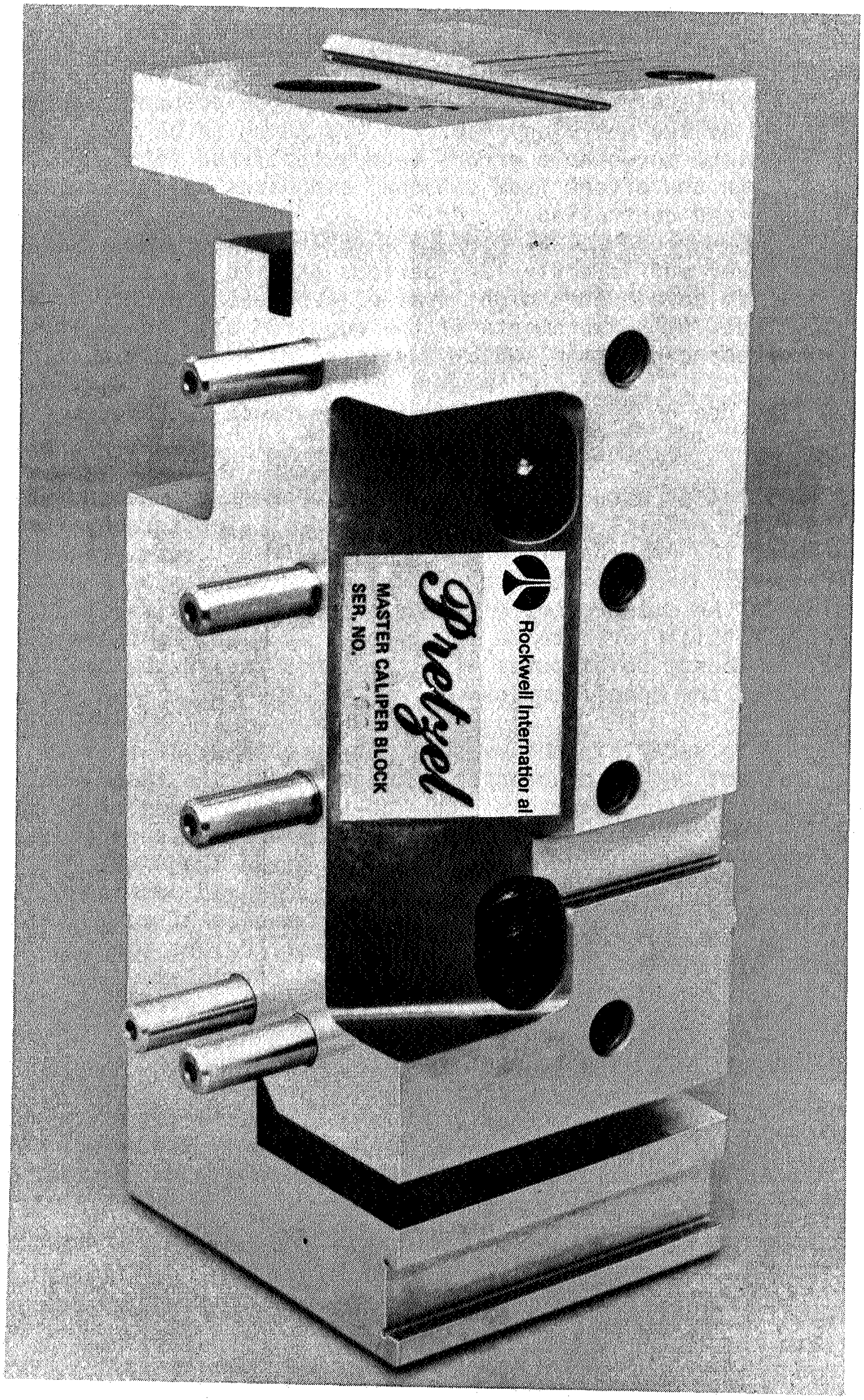

FIGURE 8. MASTER CALIPER BLOCK. (FIGURE REPRINTED WITH PERMISSION OF ROCKWELL INTERNATIONAL.)

This program is an example of the application of MAP techniques. Firstly, the measurement problem was well-defined. The company had identified a list of measurements for which they wanted to determine the ability of their machinists to generate accurate results. Secondly, the way in which the program was to operate was clearly spelled out in writing. A transport standard (the pretzel) was developed whose properties were well-established by the metrology lab. It was known to be rugged and stable. The maximum acceptable errors were established for each type of measurement. Both the offset from national standards and random error were established and controlled.

Data were kept over sufficiently long periods of time to detect shifts or trends. This firm developed a high level of assurance that their machinists were capable of making measurements of the required accuracy. It illustrates how the MAP concept can be applied at any level of a measurement echelon.

8.6 The Use of Measurement Quality Assurance Techniques in Federal Agencies

The use of measurement assurance techniques by a Federal Government agency in carrying out its mission is described in Bibliography item 19. In this example, the Army had to procure lasers from several vendors. Early in the development phase, Army metrologists became concerned that discrepancies in the measurement of laser output might exist among the contractors. If this proved to be true, the Army might be faced with costly and time-consuming disputes with the vendors over whether their laser systems met the specifications.

The Army needed to ascertain the measurement capability of their contractors and detect any discrepancies between the Army's laser measurements and those of the contractors. During the first "round-robin" it was discovered that their concerns were well-founded: discrepancies as large as 200 percent were found. The vendors shared the Army's view of the importance of achieving consistent measurements and promised their cooperation. Those vendors without effective measurement quality control programs were urged to institute such techniques. Before long, the sources of the discrepancies were identified and corrected. The maximum disagreement was reduced by a factor of 5, and ongoing efforts indicate that the consistency of data among vendors is continuing to improve. No disputes over lasers failing to meet specifications should develop as a result of measurement incompatibilities. The improvements instituted by those vendors who uncovered and corrected measurement problems will benefit their civilian customers as well. Had contract litigation developed over alleged failure to meet specifications, it is probable that the cost of engineering, management, and legal personnel required for litigation would have greatly exceeded the cost of developing the improved measurement methods for laser output.

Bibliography item 20 describes a MAP developed by a Federal Agency concerned with the enforcement of coal mine safety regulations.

8.7 The Use of Measurement Quality Assurance Techniques in State and Local Governments

The National Bureau of Standards is providing technical guidance and support to State weights and measures laboratories to develop measurement control programs. The objective is to establish a continuing program for each facility that:

- o Monitors the offset of the laboratory standards originally provided by NBS;
- o Establishes and updates the uncertainty of these State laboratories' measurements;
- o Monitors the performance of the State laboratory metrologists.

While other techniques have been in use to meet these objectives, the new control programs being put into effect greatly increase the comprehensiveness of the program with a minimum amount of additional workload for the State laboratories and for NBS. (See Bibliography item 54 for more details.)

The State laboratories typically provide calibrations and tolerance testing* in the areas of mass, volume, and length. Some laboratories provide services in other measurement areas. The bulk of their workload is in tolerance testing and calibrating mass standards. The mass calibration area demands the greatest precision and is the first area in which a measurement control program was developed and implemented.

Almost all mass calibrations performed in State laboratories are on weight sets from 1 mg to 100 g. Normally, two balances are used for this range of weights. Procedures require the metrologist to intercompare the State mass standards during the time an unknown weight set is being calibrated. Three decades (sets) of State standards are intercompared in this process; the 100 g versus summation of 50, 30, and 20 g, the 10 g versus summation of 5, 3, and 2 g, and the 100 mg versus summation of 50, 30, and 20 mg. The 100 g decade is intercompared on the larger capacity balance used in the calibration process and the 10 g and 100 mg decades are intercompared on a microbalance. These three intercomparisons serve to meet all the objectives of the control program.

In this program, the results of each intercomparison are analyzed to determine whether or not the process is in control. The measured difference between the 100 g weight and the summation of the 50 g, 30 g, and 20 g weights is checked for consistency with the values assigned to the weights by NBS. The test for agreement takes into account both the uncertainty in NBS' original assignment of values and the precision of the measurement process in the State laboratory.

If the observed difference lies within the expected limits, the measurement process is assumed to be in control, and measurements made on unknown weights are assumed to be valid within the stated uncertainty limits. If the observed difference lies outside the expected limits, the integrity of the measurement process is suspect, and measurements of

* Tolerance testing is defined in the Glossary under "calibration."

unknown weights made since the previous intercomparison are then also suspect.*

This procedure does not provide an absolute determination of systematic error. For example, if some weights used in the summation were to lose mass and others were to gain comparable mass, such a change might not be detected. However, it does provide a means for detecting common systematic errors. If a discrepancy between the current measurements and the assigned values is noted, the state laboratory's weights are returned to NBS for recalibration.

Control charts are established by intercomparing the standards over time. The variability in the data establishes the random errors present in the measurement process. As each unknown weight set is calibrated, additional data on the State standards are generated and added to the control chart. Because measurements are made throughout the year, all the parameters likely to affect the measurement process should be reflected in the control charts. New measurements must fall within the control limits on the charts for the measurements to be in control.

The control charts are also used to monitor the stability of the standards. If the standards are changing, this will be evident by a trend in the data which can be readily observed from the control chart. Continuous collection of data will reveal this drift early in its development and permit the metrologist to take corrective action before the drift causes a serious problem.

If there is more than one person performing the measurement, the data can also be used to determine if there is a difference in results dependent upon the operator of the balance. Ideally, there should not be a significant difference in results between operators, but data collected for each operator can be used to determine if this is the case.

Since this is a continuing program, the data on the standards are statistically analyzed to compare the results from previous years to the results of the current year. First, the average of the data on each decade of standards is compared using the t-test. Next, the precision of the current data is compared to that of previous data using the F-test. (See Part II for a discussion of these tests.) If the results are consistent, they are combined and a new control chart is made based upon the updated data. If problems or changes are revealed, they are investigated and corrective action taken. This establishes a continuous, comprehensive, internal measurement control program in a State laboratory.

The final aspect of the measurement control program is to verify that the results among State laboratories agree. This part of the program is accomplished through Regional Measurement Management Program (MMP) groups. These groups consist of the metrologists of the State laboratories in a particular geographic region. Currently, only three groups are operational, but two additional groups have been formed that complete coverage of the United States.

---

* Some State laboratories use weighing designs that utilize check standards in each measurement of an unknown weight to document the validity of each measurement. When the measurements are particularly critical, this additional effort can often be justified.

These MMPs include round-robin testing on NBS calibrated standards and technical meetings to discuss test methods and address regional problems. With guidance from NBS, the MMP groups develop round-robin experiments, coordinate the movement of the standards to the member labs, and analyze the data. The NBS assists in the investigation of problems.

At the present time, only the measurement control program in mass calibration has been implemented in this manner. Several other control programs in other measurement areas are being designed and will be analogous to the mass calibration program. In 1983, for example, control charts for volume transfer using glass standards will be required of the state laboratories. Simpler programs will be implemented in tolerance testing measurements where the tolerances are relatively large compared to the variability in the measurement process.

The strength of the State weights and measures laboratory measurement assurance approach is that it is applicable to a wide range of measurements and is flexible enough to permit each measurement control program to be tailored to the particular needs of a given measurement area. The sophistication of a control program varies with the criticalness of the measurement. If a measurement has relatively large tolerances and the laboratory equipment is very precise, a very simple measurement control program can be implemented. If a high degree of measurement assurance is required, a more complex measurement control system can be used. The advantage of a properly designed measurement assurance program is that a large amount of information can be obtained with a minimum amount of work.

This program minimizes the amount of NBS resources needed, and establishes measurement control among State laboratories, with the NBS as the unifying base for the country.

Bibliography

General References

1. Cameron, J. M. Measurement Assurance. Natl. Bur. Stand. (U.S.) NBSIR 77-1240; 1977 April.

2. Cameron, J. M. Measurement Assurance. Review of Standards and Specifications Section of Journal of Quality Technology 8(1): 53-55; 1976 January.

3. Pontius, P. E. Notes on the Fundamentals of Measurement and Measurement as a Production Process. Natl. Bur. Stand. (U.S.) NBSIR 74-545, 1974 September.

4. Schumacher, Rolf B. F. New Metrology Concepts Through Product QC Methods. Transactions of the 36th Annual Technical Conference of the American Society for Quality Control; 1982 May 3-5; Detroit, MI, 153-157.

Traceability

5. Belanger, B. C. Traceability: An Evolving Concept. ASTM Standardization News, 8(1): 22-28; 1980 January.

6. Cameron, Joseph M. Traceability. J. Quality Technology 7(4): 193-195; 1975 October.

Regional MAPs

7. Belanger, Brian C.; Kieffer, Lee J. Regional Measurement Assurance Programs for Physical Measurements. Proceedings of the IMEKO 8 Conference; 1979, May 21-27 Moscow, U.S.S.R. n.p. (Reprints available from the NBS Office of Physical Measurement Services.)

8. Davidson, Gary M. Regional Measurement Assurance Programs Past and Future. Transactions of the 1980 ASQC Technical Conference; 1980 April 20-22; Atlanta, Georgia, 16-23.

Statistical Concepts, Precision and Accuracy, Experimental Designs

9. Ku, Harry H., ed. Precision Measurement and Calibration Vol. 1, Selected NBS Papers on Statistical Concepts and Procedures. Natl. Bur. Stand. (U.S.) Spec. Publ. 300, Vol. 1; 1969 February. (This reference contains many landmark papers on this subject that are deserving of separate citations in their own right.)

10. Natrella, Mary G. Experimental Statistics. Natl. Bur. Stand. (U.S.) Handbook 91; 1966 January.

11. ASTM Manual on Presentation of Data and Control Chart Analysis, ASTM Special Technical Publication 15D; 1976, October (Published by the American Society for Testing and Materials, Philadelphia, Pa.)

12. Eisenhart, Churchill. Realistic Evaluation of the Precision and Accuracy of Instrument Calibration Systems. J. Res. Natl. Bur. Stand. (U.S.) 67C(2): 161-187; 1963 April-June.

13. Silverman E. N.; Brody, L. A. Statistics--A Common Sense Approach. Boston, Mass. Prindle, Weber, and Schmidt, Inc., 1973.

14. Cameron, Joseph M. The Use of the Method of Least Squares in Calibration. Natl. Bur. Stand. (U.S.) NBSIR 74-587; 1974 September.

15. Schumacher, Rolf B. F. Measurement Assurance Through Control Charts. 33rd Annual ASQC Technical Conference Transactions. 1979 May 14-16; Houston, Texas, 401-409.

16. Box, G. E. P.; Hunter, W. G.; Hunter, J. S. Statistics for Experimenters. New York: John Wiley & Sons, Inc., 1978.

17. Freedman, D.; Pisam, R.; Purves, R. Statistics. New York: W. W. Norton and Co., Inc., 1978.

Measurement Assurance Programs Outside NBS

18. Hall, J. A. Dimensional Measurement Assurance Programs. Autonetics Report No. X731146/031, Proceedings of the 1973 Conference of the National Conference of Standards Laboratories. 1973 November, Gaithersburg, Md.

19. Fecteau, M. L. The Army Laser Measurement Assurance Program: ALMAP. Proceedings of the Society of Photo-Optical Instrumentation Engineers Conference "Systems Aspects of Electro-Optics," Vol. 187. 1979 May 22-23; Huntsville, Alabama, 11-14.

20. Parobeck, P.; Tomb, T.; Ku, H. H.; Cameron, Joseph M. Measurement Assurance Program for Weighing of Respirable Coal Mine Dust Samples. J. Quality Tech. 13(3): 157-165, 1981 July.

Quality Control Methodology

21. Juran, J. M.; Gryna, F. M., Jr.; Bingham, R. S., Jr., eds. Quality Control Handbook. Third Edition, New York: McGraw-Hill Book Co. 1974.

22. Burr, I. W. Elementary Statistical Quality Control. New York: Marcel Dekker, Inc., 1979.

Specific Measurement Assurance Programs

Mass

23. Pontius, Paul E. Measurement Philosophy of the Pilot Program for Mass Calibration. Natl. Bur. Stand. (U.S.) Tech. Note 288; 1966 May.

24. Pontius, Paul E. Mass and Mass Values. Natl. Bur. Stand. (U.S.) Monogr. 133; 1974 January.

25. Cameron, Joseph M.; Croarkin, M. Carroll; Raybold, Robert C. Designs for the Calibration of Standards of Mass. Natl. Bur. Stand. (U.S.) Tech. Note 952; 1977 June.

26. Pontius, Paul E; Cameron, Joseph M. Realistic Uncertainties and the Mass Measurement Process. Natl. Bur. Stand. (U.S.) Monogr. 103; 1967 August.

27. Varner, Ruth N.; Raybold, Robert C. National Bureau of Standards Mass Calibration Computer Software, Natl. Bur. Stand. (U.S.) Tech. Note 1127; 1980 July.

DC Voltage

28. Eicke, Woodward G. Jr; Auxier, Laurel M. Regional Maintenance of the Volt Using Volt Transfer Techniques. IEEE Transactions on Inst. and Meas., IM23(4): 290-294; 1974 December.

29. Eicke, Woodward G.; Cameron, Joseph M. Designs for Surveillance of the Volt Maintained by a Small Group of Saturated Standard Cells. Natl. Bur. Stand. (U.S.) Tech. Note 430; 1967 October.

Resistance

30. Hermach, Frank L.; Dziuba, Ronald F., eds. Precision Measurement and Calibration: Electricity--Low Frequency. Natl. Bur. Stand. (U.S.) Spec. Publ. 300, Vol. 3; 1968 December (several articles on resistance measurement).

Capacitance

31. Free, George; Morrow, Jerome. Transportable 1000 pF Standard for the NBS Capacitance Measurement Assurance Program. Natl. Bur. Stand. (U.S.) Tech. Note 1162; 1982 November.

32. Levy, Charles R., Testing to Quantify the Effects of Handling of Gas Dielectric Standard Capacitors. Natl. Bur. Stand. (U.S.) Tech. Note 1161; 1982 October.

Electrical Energy (Watthour meter)

33. Oldham, N. Michael. A Measurement Assurance Program for Electric Energy. Natl. Bur. Stand. (U.S.) Tech. Note 930; 1976 September.

34. Ramboz, J. D.; McAuliff, R.C. A Calibration Service for Wattmeters and Watthour Meters. Natl. Bur. Stand. (U.S.) Tech. Note 1179; 1983 July. (For additional information, see also Houghton, S. R. Transfer of the Kilowatthour. IEEE Trans. on Power Apparatus and Systems, PAS94(4): 1232-1240; 1975 July-August.)

Temperature

35. Riddle, J. L.; Furukawa, G. T.; Plumb, H. H. Platinum Resistance Thermometry. Natl. Bur. Stand. (U.S.) Monogr. 126; 1973 April.

36. Furukawa, George T.; Riddle, J. L.; Bigge, William R. The International Practical Temperature Scale of 1968 in the Region 90.188 K to 903.89 K as Maintained at the National Bureau of Standards. J. Res. Natl. Bur. Stand. (U.S.) 80A: 477-504; 1976.

37. Furukawa, George T.; Riddle, J. L.; Bigge, William R. The International Practical Temperature Scale of 1968 in the Region 13.81K to 90.188K as Maintained at the National Bureau of Standards. J. Res. Natl. Bur. Stand. (U.S.) 77A: 309-332; 1973 May-June.

Laser Power and Energy

38. West E. D.; Churney, K. L. Theory of Isoperibol Calorimetry for Laser Power and Energy Measurements. J. Appl. Phys. 41(6): 2705-2712; 1970 May.

39. West, E. D.; Case, W. E.; Rasmussen, A. L.; Schmidt, L. B. A Reference Calorimeter for Laser Energy Measurements. J. Res. Natl. Bur. Stand. (U.S.) 76A(1): 13-26; 1972 January-February.

40. Franzen D. L.; Schmidt, L. B. Absolute Reference Calorimeter for Measuring High Power Laser Pulses. Applied Optics 15: 3115; 1976 December.

41. West E. D.; Schmidt, L. B. A System for Calibrating Laser Power Meters for the Range 5-1000 W. Natl. Bur. Stand. (U.S.) Tech. Note 685; 1977 May.

42. Danielson, B. L. Measurement Procedures for the Optical Beam Splitter Attenuation Device BA-1. Natl. Bur. Stand. (U.S.) NBSIR 77-858; 1977 May.

43. Sanders, A. A.; Rasmussen, A. L. A System for Measuring Energy and Peak Power of Low-Level 1.064 Micrometer Laser Pulses. Natl. Bur. Stand. (U.S.) Tech. Note 1058; 1982 November.

44. Case, W. E. Quality Assurance Program for the NBS C, K, and Q Laser Calibration System. Natl. Bur. Stand. (U.S.) NBSIR 79-1619; 1979 August.

45. Sanders, A. A.; Cook, A. R. An NBS Laser Measurement Assurance Program. Proceedings of the EOSO Conference Technical Program; 1976 September, 14-16; New York City, N.Y., 277-280.

Gage Blocks

46. Croarkin, M. Carroll; Beers, John; Tucker, Clyde D. Measurement Assurance for Gage Blocks. Natl. Bur. Stand. (U.S.) Monogr. 163; 1979 February.

47. Pontius, Paul E. Measurement Assurance Program--A Case Study: Length Measurements. Part 1. Long Gage Blocks (5 in. to 20 in.). Natl. Bur. Stand. (U.S.) Monogr. 149; 1975 November.

48. Beers, John S. A Gage Block Measurement Process Using Single Wavelength Interferometry. Natl. Bur. Stand. (U.S.) Monogr. 152; 1975 December.

49. Beers, John S.; Tucker, Clyde D. Gage Block Flatness and Parallelism Measurement. Natl. Bur. Stand. (U.S.) NBSIR 72-239; 1973 August.

50. Beers, John S.; Tucker, Clyde D. Intercomparison Procedures for Gage Blocks Using Electromechanical Comparators. Natl. Bur. Stand. (U.S.) NBSIR 76-979; 1976 January.

51. Beers, John S.; Taylor, James E. Contact Deformation in Gage Block Comparisons. Natl. Bur. Stand. (U.S.) Tech. Note 962; 1978 May.

52. Tucker, Clyde D. Preparations for Gage Block Comparison Measurements. Natl. Bur. Stand. (U.S.) NBSIR 74-523; 1974 July.

53. Varner, Ruth N. Computer Software for Measurement Assurance of Gage Blocks. Natl. Bur. Stand. (U.S.) Tech. Note 1168; 1982 October.

State MMP for Weights and Measures

54. Oppermann, Henry V. State Regional Measurement Assurance Programs, NCSL Newsletter 23(1): 21-26; 1983 March.

Appendix

Glossary of Terms

Universally agreed to definitions do not exist for the terms defined below. The definitions here should be regarded as interim definitions proposed by the author. The reader should refer to Part II for a more detailed discussion of statistical terminology.

Accepted Value (Consensus Value) The value assigned to a standard or instrument in a calibration process, the validity of which is accepted explicitly or implicitly by the parties affected by the calibration. For many types of measurements in the United States, the accepted value would be the value assigned to a standard when it is calibrated by NBS. (It should be noted that the value assigned to a standard calibrated by NBS is considered by NBS to be valid only at the time of calibration.) In the context of measurement process parameters as described in Part II, "accepted value" refers to a historically-determined value of the process parameter.

Accuracy: The extent to which the measured value of a quantity agrees with the accepted or consensus value for that quantity. (See Chapter 3 for a discussion of what is meant by accepted or consensus value.)

Calibration: Calibration can be defined in two senses:

1. The process of assigning values to the response of an instrument or the property of an artifact relative to a reference base.

2. The comparison of a measurement system or device of unverified accuracy to a measurement system or device of known and greater accuracy to detect or correct any variation from required performance specifications of the unverified measurement system or device.

   It should be noted that in some organizations "calibration" includes only the determination of whether or not the particular instrument or standard is within some established tolerance. (Called "tolerance testing" in Section 8.7.) In other organizations "calibration" includes the reporting of deviations from nominal values. In still other organizations "calibration" additionally includes any repair/adjustment required to bring the item back into the established tolerance.

<u>Check Standard</u>: A stable, well-characterized in-house standard that is remeasured at periodic intervals to determine whether the measurement process is in a state of statistical control. For some measurement processes some combination of measurements is used to monitor statistical control (such as the difference in length between two nominally equal gage blocks). In such cases this control parameter may also be referred to as a "check standard".

<u>Consensus Value</u> See Accepted Value

<u>Control Chart</u>: A graphical tool for ascertaining whether or not a measurement process is or is not in a state of statistical control. (See Part II.)

<u>Limiting Mean</u>: The value approached by the average of a sequence of independent measurements of the same quantity as the number of measurements included in the set approaches infinity.

<u>Measurement Assurance Program (MAP)</u>: A quality assurance program for a measurement process that quantifies the total uncertainty of the measurements (both random error and systematic components of error) with respect to national or other designated standards, and demonstrates that the total uncertainty is sufficiently small to meet the user's requirements.

<u>Measurement Process</u>: A sequence of operations whose purpose is to assign a number (or numbers) which represent(s) how much of a certain property a given substance or object has, or how a certain property of a substance or object relates to other properties of the same or other substances or objects.

<u>Offset</u>: See systematic error.

<u>Pivot Laboratory</u>: A laboratory that plays a coordinating role on behalf of a group of laboratories participating in a regional or group MAP.

The role of the pivot laboratory may vary somewhat from one MAP to another; it may include one or more of the following functions:

- Coordinating the movement of the transport standard among NBS and the group member laboratories,
- Analyzing measurement data from the group,
- Disseminating MAP information from NBS and providing feedback to NBS,

- Serving as a site to which other group MAP participants can bring their standards for calibration and/or intercomparison.

In some MAPs the pivot laboratory is permanently designated as such. In other cases, laboratories participating in the group MAP take turns being the pivot laboratory.

Precision: The degree of agreement among independent measurements of a quantity under specified conditions.

For a set of data, the standard deviation (or some multiple thereof) is frequently taken as a measure of precision (more correctly, the standard deviation is a measure of the imprecision). A process may have high precision but still be inaccurate. (This would be the case where the random error is small but the systematic error or offset is large.)

Quality Assurance: All those planned or systematic actions necessary to provide adequate confidence that a product or service will satisfy given needs.

Quality Control: The operational techniques and the activities which sustain a quality of product or service that will satisfy given needs; also the use of such techniques and activities.

Random error: Random error is an error which varies in an unpredictable manner in absolute value and in sign when measurements of the same value of a quantity are made under effectively identical conditions.

Repeatability: The closeness of the agreement among the results of successive measurements of the same quantity carried out by the same method, by the same observer, with the same measuring instruments, in the same laboratory, at quite short intervals of time.

Reproducibility: The closeness of the agreement among the results of measurements of the same quantity, where the individual measurements are made:

- By different methods, with different measuring instruments
- By different observers, in different laboratories
- After intervals of time quite long compared with the duration of a single measurement
- Under different normal conditions of use of the instruments employed.

Round-robin: In the context of this publication, a systematic intercomparison among laboratories or other measurement locations wherein a transport standard, set of standards, or other artifact is circulated for measurement among the participants to evaluate the offset of each participant's measurements on the artifact relative to the other participant's measurements and to the accepted value of the property being measured.

State of Statistical Control (of a Measurement Process):
Lay person's definition: A measurement process is in a state of statistical control if the amount of scatter in the data from repeated measurements of the same item over a period of time does not change with time and if there are no sudden shifts or drift in the data.

Statistician's Definition: A measurement process is in a state of statistical control if the resulting observations from the process, when collected under any fixed experimental conditions within the scope of the a priori well-defined conditions of the measurement process, behave like random drawings from some fixed distribution with fixed location and fixed scale parameters.

Systematic error (offset):
An error which, in the course of a number of measurements, made under the same conditions, of the same value of a given quantity, either remains constant in absolute value and sign, or varies according to a definite law when the conditions change. The systematic error of a measurement process is the difference between the limiting mean of independent measurements of the measured quantity and the true value of that quantity. Since the true value is generally unknown, the difference between the limiting mean and the "accepted value" is taken to be the systematic error.

In a calibration hierarchy, the random error obtained by NBS when comparing a laboratory's standard to a national reference standard should be considered to be a systematic error when the standard is used for calibrations by the laboratory.

Transfer Standard (see also transport standard):
Any standard that is used to intercompare a measurement process at one location or level with that at another location or level. The term is often used interchangeably with the term transport standard.

Transport Standard (see also transfer standard):
A rugged, well-characterized transportable standard that can be packaged for shipping back and forth between NBS and the MAP participants. In a MAP, NBS usually measures the transport standard before and after it is measured by the participant(s).

True Value: The correct or actual value of the quantity being measured.

(See "accepted value" and the discussion in Chapter 3.) The true value of a reference artifact will often vary with time. For example, the true length of a reference end bar or gage block changes very slightly from instant to instant because the gage block expands and contracts as a result of minute temperature fluctuations.

Uncertainty: The maximum credible limits for the difference between the accepted or consensus value and the measured value of the quantity of interest (or between the true value and the measured value). These limits may be unsymmetrical in the most general case and represent the sum of a measure of the random error and estimated bounds to the systematic error(s). Uncertainty, therefore, is a quantification of inaccuracy and imprecision. If a reported value is said to have an uncertainty of one part per million, this means that the person reporting the uncertainty has evidence to demonstrate that the reported value should not differ from the true value or consensus value by more than one part per million.

*U.S. GOVERNMENT PRINTING OFFICE : 1985 0-461-105/20053