A11106 082652

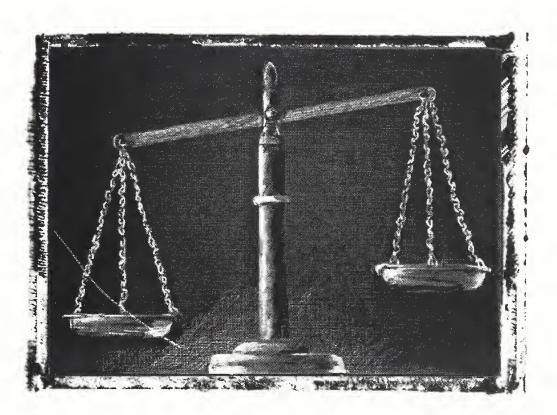
NIST

REFERENCE

NISTIR 6855

Overview of the Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales

Val R. Miller



contribution of the National Institute of Standards and Technology; not subject to copyright in States."

QC 100 .U56 #6855 2002

NISTIR 6855

Overview of the Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales

Val R. Miller Office of Weights and Measures

January 2002



U.S. Department of Commerce Donald L. Evans, Secretary

Technology Administration Phillip J. Bond, Under Secretary for Technology

National Institute of Standards and Technology Arden L. Bement, Jr., Director





Overview of the Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales

Abstract:

As scale and balance service organizations have become ISO 9000 registered or accredited to ISO Guide 25 and ANSI/NCSL Z540-1-1994, now implemented as ISO/IEC 17025, the number of requests for assistance in calculating the uncertainty of a scale or balance calibration have dramatically increased. It has become evident that a practical guide is needed to explain the calculation of uncertainty in a manner usable by field service personnel who are not trained statisticians.

ISO/IEC 17025 states that a calibration laboratory or testing laboratory performing calibrations shall have and shall apply a procedure to estimate the uncertainty of measurement for all calibrations. This guide will attempt to clarify the concepts of calculating scale or balance calibration uncertainty, by providing specific examples, and making the interpretation of the various methodologies understandable to field personnel.

Key words: balance; GUM; ISO 17025; repeatability; scale; uncertainty;



Introduction:

The ISO/IEC "Guide to the Expression of Uncertainty in Measurement" (GUM) defines uncertainty as:

Uncertainty of measurement (VIM 3.9): parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand.

Notes:

- 1) The parameter may be, for example, a standard deviation (or a given multiple of it), or a half-width of an interval having a stated level of confidence.
- 2) Uncertainty of measurement comprises, in general, many components. Some of these components may be evaluated from the statistical distribution of the results of series of measurements and can be characterized by experimental standard deviations. The other components, which can also be characterized by standard deviation, are evaluated from assumed probability distributions based on experience or other information.
- 3) It is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion.

The American Heritage® Dictionary of the English Language, Fourth Edition defines uncertainty as:

Uncertainty: *"Statistical:* The estimated amount or percentage by which an observed or calculated value may differ from the true value.

SYNONYMS: doubt, dubiety, skepticism, suspicion, mistrust."

Users of scales and balances are able to make an estimate of the true value of a measured load because of the calibration process. By applying a mass artifact with a known value, the calibrator assigns to each scale or balance division a corresponding value. When an item with an unknown weight value is placed on a calibrated weighing system the calibrated division-to-weight ratio is used to estimate the value of the measured item. However, that value is an estimate; around the estimated value is a band of potential error, the measurement uncertainty. This might also be viewed as a region of doubt, mistrust, suspicion or skepticism. This paper describes the "Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales" and discusses the reasons why such a guide is necessary. It also provides background about properly including several common uncertainty issues in the uncertainty calculations.

Why write a Recommended Guide?

This field guide is needed to explain how to calculate scale and balance calibration uncertainty in a manner usable by field service personnel. The NIST Office of Weights and Measures has had a marked increase in the number of requests for assistance in calculating scale and balance uncertainties. This increase has resulted from the large number of scale and balance service organizations dealing with issues involving ISO 9000 registration or accreditation to ISO/IEC 17025.

ISO/IEC 17025 requires that measurements be traceable to national standards. A significant part of traceability is a properly documented uncertainty statement. However, most field service personnel have little training in calculating measurement uncertainty, and no specific guidance has been available regarding how to calculate and report uncertainties using terminology understood by field personnel.

ISO/IEC 17025 states that a calibration or testing laboratory performing calibrations shall apply a procedure to estimate the uncertainty of measurement for all calibrations. One method previously used to calculate calibration uncertainties was to simply restate the uncertainty of the standard mass calibrations. Another method involved stating the tolerance of the mass standards used in the test. These and other methods have been employed, but most of these methods ignored significant uncertainty contributors. When these more significant uncertainty contributors are included in the calculations the contribution of the tolerance or uncertainty of the standards becomes nearly indiscernible.

The ISO "*Guide to the Expression of Uncertainty in Measurement*" provides excellent guidance for calculating uncertainties, and describes a method to evaluate the uncertainty of a measurement process. However, it is a general guideline and is most easily understood by statisticians, not by field service technicians. Few scale and balance service providers employ a statistician to calculate the uncertainty of their measurements; therefore, few service organizations have been properly calculating and reporting scale and balance calibration uncertainties.

The complexity of the weighing system calibration uncertainty has not been understood or appreciated by most service providers. This is largely because the scale or balance test procedures, such as those in NIST Handbook 44, "*Specifications, Tolerances, and Other Technical Requirements for Weighing and Measuring Devices*"; American Society for Testing and Materials (ASTM) Standard Practices: E 319-85, "*Standard Practice for the Evaluation of Single-Pan Mechanical Balances*"; E 898-88, "*Standard Method of Testing Top-Loading, Direct-Reading Laboratory Scales and Balances*" and E 1270-88, "*Standard Test Method for Equal Arm Balances*"; and Organisation Internationale de Métrolegie Légale (OIML), R 76-1, "*Nonautomatic weighing instruments, Part 1: Metrological and technical requirements-Tests*," do not address the calculation of calibration uncertainty.

It has been the responsibility of the servicing organization to address the concept of calibration uncertainty, and most service organizations have not understood this concept. The Guide will provide guidance that scale service companies can apply to calculating the uncertainty of their calibration processes and appropriately reporting the uncertainty in their calibration reports and other documentation.

Who will benefit?

Over the last 25 years the author has read numerous articles and papers that have dealt with the concepts of uncertainty with a wide variety of approaches. None have addressed in their entirety the issues that the scale and balance service technician confronts daily. Each reference document has contained conceptual information and examples of uncertainty calculations, but no single document has expressly addressed the calibration of a scale or balance and the calculation of the uncertainty associated with such a calibration. The Guide will benefit all scale and balance service personnel who are required to provide an estimate of uncertainty for the calibrations they perform. Information from a number of documents has been interpreted and combined in a format and procedure that can be applied each time a calibration is performed.

What approach is used in this Guide to determine Balance and Scale Uncertainty?

This Guide consists of two basic sections. The first section describes an eight-step process for calculating uncertainty similar to that described in the GUM. Background information of the eight-step process and a discussion of general issues are included in this section.

The eight-step process:

Step 1: Specify the measurement process and equation

Step 2: Identify and characterize measurement influences

Step 3: Quantify uncertainty estimates

Step 4: Convert factors to standard uncertainty values

Step 5: Calculate the combined standard uncertainty

Step 6: Calculate the expanded uncertainty

Step 7: Evaluate the expanded uncertainty

Step 8: Report the uncertainty

The second section of the Guide is comprised of a series of appendices that describe how the eight-step approach is used in specific scale or balance calibration processes. The appendix format was selected to simplify the addition of new uncertainty analysis processes as they are developed. Currently, Appendix A addresses Laboratory Balance Calibrations, Appendix B addresses Large Capacity Scale Calibrations, and Appendix C covers General Purpose Scale Calibrations. Each appendix has sections that deal with individual uncertainty contributors that can be minimized either by repair of the device or by minimizing or eliminating the factors and influences that make the contributor significant. The appendices describe options for determining an uncertainty influence value and how that value should be treated for inclusion in the final uncertainty value. Each appendix also includes an Uncertainty Worksheet to assist the user in properly calculating uncertainties that may be used as part of routine uncertainty documentation.

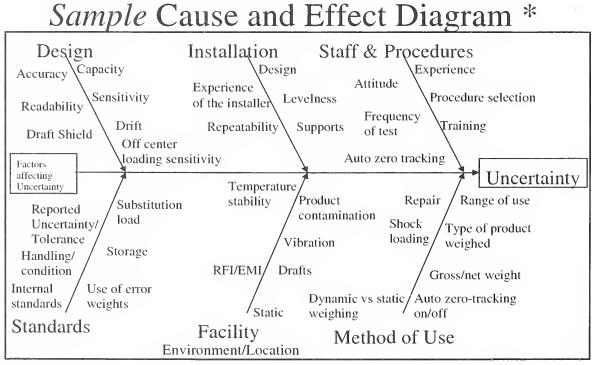
As explained in the Guide, service technicians who follow this eight-step approach will be able to determine and report a meaningful estimate of the uncertainty in the calibration performed. The following paragraphs provide insight into the structure of the Guide and describe how the Guide addresses each of the eight steps to determining measurement uncertainty.

Step 1: Specify the measurement process and equation

Without a well-defined process equation, the associated uncertainties cannot be defined. The measurement process must be defined and documented before an understanding of the measurement uncertainty influences can be developed. The methods to be employed in making the measurement must be clear and an equation developed that reflects the processes. It must be clearly understood that scale and balance calibrations do not measure the weight of the artifacts applied, but instead measure the weighing device's capability to correctly represent the weight of the applied mass artifacts.

Step 2: Identify and characterize measurement influences

An effective uncertainty tool described in the Guide is a cause and effect diagram which used to aid in identifying possible measurement influences that may contribute to the calibration uncertainty. This diagram results from a brainstorming session where all known possible uncertainty contributors are recorded in a formal manner. The six major headings of Design, Installation, Staff and Procedures, Facility, Standards and Method of use are common to most measurement uncertainty evaluations. Each of these major headings may top a virtually unlimited list of measurement influences. The examples shown in the following diagram are among the most common and significant, but remember, each situation must be evaluated independently.



* Not all inclusive

Step 3: Quantify uncertainty estimates

Each measurement influence is evaluated to determine the magnitude of the impact on the measurement. Decisions must be made about how the measurement influence will be treated before it can be included in the measurement uncertainty. These decisions will be based on the tests performed.

Many of the factors listed in the cause and effect diagram above will be considered insignificant if appropriate corrective measures are taken to minimize their impact on the measurement. Corrective measures may be made by repairing the weighing device or by minimizing the source of the error. The expected contribution of each influence factor must be documented to show that it was considered and if it was determined to be insignificant. The availability of computer spreadsheet calculations makes it possible to include the measurement impact of every known influence though individual contributions may be insignificant. Those that are insignificant will have little or no effect on the final result, but the documentation showing that the influence was evaluated will be automatically generated by the software. Without a computer spreadsheet, the service personnel must generate separate documentation to show that each influence was evaluated and whether or not it was deemed significant.

Step 4: Convert factors to standard uncertainty values

Several of the listed influence factors cannot be eliminated and must be included in all weighing system uncertainty calculations. The contribution of each measurement influence factor to the total uncertainty of the measurements must be converted to a standard deviation or a 'standard uncertainty' value. The following paragraphs show how the standard uncertainty for some uncertainty components can be estimated based on knowledge of the weighing device.

The standard deviation or repeatability of a device will always be a part of the uncertainty, as will the uncertainty of the mass standards. Large capacity truck scale calibrations performed outdoors may also be influenced by wind and varying temperatures that affect the scale indications. These influences must be evaluated to determine their significance. If it is decided that the impact on the measurement is significant, a component must be included in the uncertainty calculations for these influence factors. Using information known about the possible measurement error and the weighing device, decisions about how each of these components will be included must then be made.

Repeatability Error

The repeatability errors for balances and scales are best determined by making numerous repeated measurements of a single load. Statisticians prefer to see a minimum of thirty measurements, though practicality sometimes dictates using as few as seven to ten measurements. The standard deviation of the recorded measurement results is then calculated. The result is the repeatability error of the device (s_p) , which is included in the uncertainty calculations as a value from a

normal distribution. Occasionally, some devices with limited resolution, e.g., 50 lb scale divisions, may generate exactly the same indication for all of the repeated measurements. From a statistical point of view it is unlikely that the true repeatability error is zero, so an estimate of the repeatability error must be used. The repeatability error must be estimated from the display resolution. It is probable that the standard deviation of a large number of measurements will be a significant portion of a display digit unless a device has been designed so that the internal resolution is one or more decades greater than is displayed. (In those cases, consultations with the manufacturer are required to determine the true resolution of the display device.) Once the value of the display resolution has been determined, the standard deviation is estimated by one of two equations.

While a system standard deviation calculated from measured data is most desirable, there may be financial or time constraints that make this approach impractical, particularly when calibrating large capacity scales in an uncontrolled environment (for example, a 50 000 kg scale, in an industrial environment). In those situations, the estimated standard deviation equations described below will be applied.

If a discrimination test is not performed, it is unknown where within the range of a scale's display division (d) the true value of the applied load lies. (A discrimination test verifies a weighing system's capability to discern one value from another differing by an amount less than one half of a scale division, d). Thus, it is just as probable that the true value lies nearer to the indicated load plus one division, as that it lies nearer to the indicated load minus one division. It is only certain that the true value is somewhere within ± 1 division of the indicated value, and that any value within that range is equally probable. This is a uniform probability distribution, sometimes called a rectangular distribution. The

estimated standard deviation equals $s_p = \frac{d}{\sqrt{3}} = 0.577d$ when a discrimination test

is not performed.

If a discrimination test is performed, it is known that the weighing system is capable of sensing a change in load less than $\frac{1}{2}$ d. Therefore, it is also known that the true value is within $\frac{1}{2}$ d of the indicated value. Because the interval of uncertainty has been decreased by a factor of two, the estimated standard

deviation equals
$$s_p = 0.5 \frac{d}{\sqrt{3}} = 0.29 d$$
.

Errors due to Environment

The measurement influences due to air currents and temperature effects are similarly treated as having a uniform probability distribution.

Air Currents

The servicing technician will often observe that air currents are causing a scale indication to randomly vary over a given indication range, so will intuitively estimate the true indication based on his observations. That estimated value will

have an uncertainty associated with it. This is estimated as $\frac{1}{\sqrt{3}}$, or 0.577, times

the maximum observed deviation from the estimated true value. Since excessive estimating is being performed in this process, it is recommended that scale and balance calibrations be performed when air currents are not a factor. But as reality sometimes requires that calibrations be performed under less than ideal conditions, a means of estimating the impact on the measurement is provided.

Temperature

Varying environmental temperatures are also a reality and can contribute to measurement uncertainty. Often, calibrations of large capacity scales span several hours during which time the temperature may vary significantly. The effect on the measurement due to temperature variations can be estimated by calculating the possible measurement error. This is accomplished by multiplying the change in temperature by the temperature coefficient of the weighing system. Because this value is derived from a manufacturer's specification, the probability distribution is

considered uniform and the uncertainty is estimated as $\frac{1}{\sqrt{3}}$, or 0.577, times the

calculated potential measurement error. Calibrations involving varying temperatures are not ideal, but reality dictates that they must be performed. The described process can estimate a value for the measurement uncertainty caused by unstable temperatures.

Other Factors

The measurement uncertainty caused by other influences can similarly be estimated by the methods described above. Most uncertainty contributors, unless calculated from a significant number (more than ten readings) of measured values, will be treated as having a uniform or rectangular probability distribution and the uncertainty

for these influences is calculated as $\frac{1}{\sqrt{3}}$, or 0.577, times the calculated potential

measurement error.

Step 5: Calculate the combined standard uncertainty

When all of the standard uncertainty values have been accumulated they are combined by the Root-Summed-Squared method to arrive at a combined standard uncertainty.

Step 6: Calculate the expanded uncertainty

The combined standard uncertainty is multiplied by a coverage factor (k) appropriate for

the confidence interval desired. By convention, as defined in NIST Technical Note 1297, "*Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*," a coverage factor of k = 2 is used for a confidence interval of approximately 95 %. The use of k values other than k = 2 requires an explanation for the deviation from convention.

Step 7: Evaluate the expanded uncertainty

At this point, the calculated uncertainty should be evaluated for correctness and appropriateness using the following questions.

Have all the uncertainty contributors been evaluated and included?

Have any arithmetical errors been made?

Is the value appropriate for the device and process?

What is the greatest uncertainty contributor and how can it be reduced?

Example: Wind influences during outdoor calibrations can cause an uncertainty component of two or three scale divisions, which may be a very significant uncertainty contributor. Performing the calibration on a calmer day would reduce the wind influence component to nearly zero. Would the uncertainty then be acceptable? How can other uncertainty influences be reduced? This is a good time to evaluate possible process improvements that can be made to reduce the uncertainty. Each uncertainty component should be evaluated relative to the combined uncertainty. The measurement influences generating the greatest uncertainty component should be examined, searching for a way to reduce the measurement impact. Where possible, corrective measures should be employed to reduce the effect of each influence. Measures such as draft shields or rescheduling the calibration to a more environmentally stable day can drastically reduce the measurement uncertainty. This evaluation/corrective measures process should continue with the next largest uncertainty component until the expanded uncertainty is acceptable or is no longer being significantly reduced by the process changes. This evaluation process provides the service technician with a means for measuring process improvement.

Step 8: Reporting uncertainty

The Guide outlines requirements and provides examples for presenting the calculated scale or balance calibration uncertainty to the user. The presentation of uncertainty is no longer a simple matter of stating that "The uncertainty is \pm 'x' grams." The GUM requires a statement explaining what uncertainty components were included and some rationale of why and how they were included.

Where will the Guide be found?

The Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales will be made available to scale and balance service organizations via the Internet on the NIST Office of Weights and Measures web site (<u>www.nist.gov/owm</u>) and by hardcopy through the Government Printing Office (GPO). It will also be made available for publication in trade journals, by the National Conference of Calibration Laboratories International, and by other quality-oriented organizations. It is desired that the Guide gain maximum distribution. The NIST Office of Weights and Measures believes that the number of calls handled by its staff to date has come from just a small portion of those organizations desiring this type of guidance. Such guidance must be readily available for every scale or balance service technician to apply in making measurements.

How will the Guide be updated with new analyses?

The appendix format was chosen specifically so that new analyses can be added with a minimum effort. Any individual or organization wishing to add an appendix for a specific scale type or process is invited to submit a copy of its analysis to the NIST Office of Weights and Measures for approval. Approved submittals will be posted with the original document on the NIST OWM web site and will be made available by hardcopy upon request. This Recommended Guide will be a continuous work-in-progress and should benefit all scale and balance service organizations.

Summary

This paper has described the reason for generating *The Recommended Guide for Determining and Reporting Uncertainties for Balances and Scales*. It has provided insight into the thought processes that have resulted in the Guide, and it has described the need to make the information contained therein available to the maximum number of users. When calibrations are performed, an estimate of the true value of the applied load is the result. Around that estimate is an area of unknown error or uncertainty. It is the goal of the author to provide the scale or balance technician with an understandable method that will allow the individual to realistically estimate the magnitude of that unknown error -- the calibration uncertainty.

Definitions:

Actual scale division (d) (OIML R 76-1, edition 1992(E)):

Value expressed in units of mass of:

the difference between two consecutive indicated values, for digital indication, and the difference between the values corresponding to two consecutive scale marks, for analogue indication.

Combined standard uncertainty (GUM):

Standard uncertainty of the result of a measurement when that result is obtained from the values of a number of other quantities, equal to the positive square root of a sum of terms, the terms being the variances or covariances of these other quantities weighted according to how the measurement result varies with changes in the quantities.

Coverage factor (k factor) (GUM):

Numerical factor used as a multiplier of the combined standard uncertainty in order to obtain an expanded uncertainty

Note: a coverage factor, k, is typically in the range of 2 to 3.

(NIST Technical Note 1297: By convention k=2 is the standard coverage factor to be used. It defines a confidence interval of approximately 95%. The use of other coverage factors requires justification of the reason for use.)

Discrimination (OIML R 76-1, edition 1992(E)):

Ability of an instrument to react to small variations of load. The discrimination threshold, for a given load, is the value of the smallest additional load that, when gently deposited on of removed from the load receptor, causes a perceptible change in the indication.

Expanded uncertainty (GUM):

Quantity defining an interval about the result of a measurement that may be expected to encompass a large fraction of the distribution of values that could reasonably be attributed to the measurand

Notes:

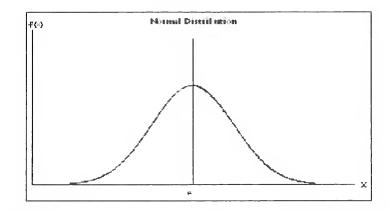
1: The fraction may be viewed as the coverage probability or level of confidence of the interval.

2: To associate a specific level of confidence with the interval defined by the expanded uncertainty requires explicit or implicit assumptions regarding the probability distribution characterized by the measurement result and its combined standard uncertainty. The level of confidence that may be attributed to this interval can be known only to the extent to which such assumptions may be justified.

3: Expanded uncertainty is termed overall uncertainty in paragraph 5 of Recommendation INC-1 (1980).

Normal distribution (paraphrased from the on-line NIST Engineering Statistical Handbook):

Distribution function for which one value (the mean) has a greater probability of occurrence than all other values. A normal distribution can be estimated with the sample mean and sample standard deviation.



The normal distribution is commonly described as being a bell shaped distribution. Note that the mean, μ , is the most probable value.

Repeatability (VIM 5.27):

Closeness of agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement.

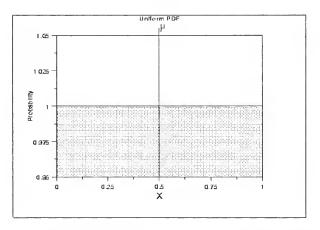
Note: 'Same conditions' includes the same procedure, same observer, same measuring instrument, same location, with repetition over a short period of time.

Standard Uncertainty (GUM):

Uncertainty of the result of a measurement expressed as one standard deviation.

Uniform distribution (paraphrased from the on-line NIST Engineering Statistical Handbook):

Distribution function where any value within the range of the function is equally probable.



Sometimes called a rectangular distribution due to its rectangular shape. This distribution function is considered the most conservative estimate of uncertainty, i.e. it gives the largest standard deviation.

References:

International Organization for Standardization, "Guide to the Expression of Uncertainty in Measurement", (GUM), International Organization for Standardization, 1993

Organisation Internationale de Métrolegie Légale (OIML), OIML R 76-1, "Nonautomatic weighing instruments, Part 1: Metrological and technical requirements-Tests", OIML, Edition 1992(E)

National Institute of Standards and Technology Technical Note 1297, 1994 Edition, "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results", National Institute of Standards and Technology, 1994

Eurachem Working Group on Chemical Measurement, "Quantifying Uncertainty in Analytical Measurement", First Edition, 1995, English edition, Crown Copyright

National Bureau of Standards, Handbook 145, SOP 29, Standard Operating Procedure for the Assignment of Uncertainty, U.S. Department of Commerce/National Bureau of Standards, 1986

Houghton Mifflin Company, The American Heritage® Dictionary of the English Language, Fourth Edition., 2000, Houghton Mifflin Company

Carroll Croarkin and Paul Tobias, "NIST/SEMATECH Engineering Statistics Internet Handbook" (<u>http://www.itl.nist.gov/div898/handbook/</u>), National Institute of Standards and Technology

Bibliography:

International Organization for Standardization, "Guide to the Expression of Uncertainty in Measurement", International Organization for Standardization, 1993 (<u>http://www.iso.ch/iso/en/ISOOnline.frontpage</u>)

Organisation Internationale de Métrolegie Légale (OIML), OIML R 76-1, "Nonautomatic weighing instruments, Part 1: Metrological and technical requirements-Tests", OIML, Edition 1992(E) (<u>http://www.oiml.org/</u>)

National Institute of Standards and Technology Technical Note 1297, 1994 Edition, "Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results", National Institute of Standards and Technology, 1994

Eurachem Working Group on Chemical Measurement, "Quantifying Uncertainty in Analytical Measurement", First Edition, 1995, English edition, Crown Copyright (<u>http://www.eurachem.bam.de/)</u>

National Bureau of Standards, Handbook 145, SOP 29, Standard Operating Procedure for the Assignment of Uncertainty, U.S. Department of Commerce/National Bureau of Standards, 1986

Carroll Croarkin and Paul Tobias, "NIST/SEMATECH Engineering Statistics Internet Handbook" (<u>http://www.itl.nist.gov/div898/handbook/</u>), National Institute of Standards and Technology

Georgia Harris, "Ensuring Accuracy and Traceability of Weighing Instruments", ASTM Standardization News, 21(4): 44-51, 1993

American Society for Testing and Materials, E 319-85 (Reapproved 1997), "Standard Practice for the Evaluation of Single-Pan Mechanical Balances", Annual Book of ASTM Standards, Vol 14.02, (<u>http://www.astm.org/</u>)

American Society for Testing and Materials, E 898-88 (Reapproved 2000), "Standard Method of Testing Top-Loading, Direct-Reading Laboraotry Scales and Balances", Annual Book of ASTM Standards, Vol 14.02

American Society for Testing and Materials, E 1270-88 (Reapproved 1997), "Standard Test Method for Equal Arm Balances", Annual Book of ASTM Standards, Vol 14.02

Walter E. Kupper, Dr.scn.techn, "High Accuracy Mass Measurement, From Micrograms to Tons", ISA Transactions (Vol 29,4,1990)

Deutscher Kalibrierdienst, DKD-R-7-1, Issue 98, "Calibration of Non-Automatic Electronic Weighing Instruments", Deutscher Kalibrierdienst, Nov. 1998





