SEMI E89-1106[E]
GUIDE FOR MEASUREMENT SYSTEM ANALYSIS (MSA)

This standard was technically approved by the global Metrics Committee. This edition was approved for publication by the global Audits and Reviews Subcommittee on May 16, 2006. It was available at www.semi.org in October 2006 and on CD-ROM in November 2006. Originally published September 1999; previously published March 2006.

[E] This standard was editorially modified in August 2006 to correct an error. Changes were made to ¶ 10.3, Note 25, and Equation 28.

SUPERSEDED

1 Purpose

1.1 The purpose of this guide is to provide a consistent set of terminology and describe a simplified, but constructive, experimental approach to planning and performing a measurement system analysis (MSA).

1.2 The goal of an MSA is to characterize the performance capability of the measurement system (MS) as it is intended to be used in a manufacturing or laboratory setting.

1.2.1 Accurately identifying the MS bias and the size and nature of all sources of variability allows one to determine whether the MS is capable of performing its intended function. Moreover, a well-designed MSA can be used to identify and quantify areas that need the most improvement.

2 Scope

2.1 This guide covers procedures for determining specific measures of MS capability including:

- measurement variability (i.e., reproducibility) under a variety of conditions, including
  - effects of repeatability,
  - load-unload, and
  - time, and
- bias, including bias-related
  - linearity,
  - stability, and
  - matching tolerance.

2.2 This guide also covers secondary metrics such as precision-over-tolerance (P/T) ratio and signal-to-noise ratio (SNR).

2.3 The primary focus of this guide is on determining measurement capability of automated wafer MSs under normal operating conditions, but the definitions and methodologies are extendible to many other measurement situations involving automated measurements on units such as processed dice, packaged devices, flat panel displays, piece parts, etc.

2.4 While there is no universally accepted correct way to conduct an MSA, the approach described in this document is supported in the technical literature (see § 11) and congruent with practices advocated in ISO 5725-2. The procedures given in this guide represent an approach to the conduct of an MSA and provide basic reference methods that should serve for a variety of applications. Other methods may be appropriate in certain circumstances.

2.5 The procedures in this guide that are intended to separate the various sources of nonsystematic (i.e., random) errors are based on the use of factorial experiments and analysis of variance (ANOVA). Because the primary focus of this guide is on evaluation of automated MSs, the variability introduced by different operators is expected to be minimal.

NOTE 1: Information on measurement uncertainty calculations is provided in Related Information 1. Information on testing measurement distributions for normality and equal repeatability is provided in Related Information 2.

NOTICE: This standard does not purport to address safety issues, if any, associated with its use. It is the responsibility of the users of this standard to establish appropriate safety and health practices and determine the applicability of regulatory or other limitations prior to use.

Documents Document Sharing Hub

Full document details: SEMI E89-1106E

3 Limitations

3.1 Determination of MS capability is meaningless unless the MS is in control. Methodology for establishing and maintaining MS control is beyond the scope of this guide. Such methodology should be a part of a quality management system, such as that mandated by ISO 9000 or similar standards. Additional guidance for laboratories without established procedures may be found in the ASTM *Manual on Presentation of Data and Control Chart Analysis.*[1]

3.2 This guide does not address those aspects of measurement uncertainty associated with change in the object being measured, either spatially or temporally.

3.3 This guide does not address determination of measurement capability in the case of destructive measurements on samples, or when the MS alters the object being measured as a result of making the measurement.

3.4 This guide does not apply to inter-laboratory experiments designed to measure inter-laboratory precision of test methods.

4 Referenced Standards and Documents

4.1 *ISO Standards*[2]

ANSI/ISO Z540-2 — Guide to the Expression of Uncertainty in Measurement

ANSI/ISO/ASQC A3534-1 — Statistics – Vocabulary and Symbols – Part 1: Probability and General Statistical Terms

ISO 3534-3 — Statistics – Vocabulary and Symbols – Part 3: Design of Experiments

ISO 5725-2 — Accuracy (trueness and precision) of measurement methods and results – Part 2: Basic method for the determination of repeatability and reproducibility of a standard measurement method

ISO 9000 — Quality management systems – Fundamentals and vocabulary, 2000

NOTICE: Unless otherwise indicated, all documents cited shall be the latest published versions.

5 Terminology

5.1 Terminology in this section that is not directly used in this guide, is likely to be encountered while conducting an MSA.

5.2 Definitions of many other terms related to metrology and statistics can be found in VIM,[3] ANSI/ISO/ASQC A3534-1, and ISO 3534-3.

5.3 *Definitions*

5.3.1 *accuracy* — closeness of agreement between a test result or the mean of a group of test results made on an object and its true value.[3]

5.3.1.1 *Discussion* — Accuracy depends on both the precision and bias of the measurement process. Since random components of error (resulting in imprecision) and systematic components of error (resulting in bias) cannot be completely separated in routine use, the reported accuracy must be interpreted as a combination of these two elements.

5.3.2 *bias* — difference between the population mean of the test results from a measurement process and the true (accepted reference) value of the property being measured.

5.3.2.1 *Discussion* — Bias is a systematic component of measurement uncertainty. One or more systematic error components may contribute to the bias. The true value and the population mean are both unknown. The true value may be estimated with the use of a consensus value. If sufficient measurements are made to adequately mitigate the effects of measurement variability, the population mean may be estimated from the sample mean

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1}$$

where:

1 *Manual on Presentation of Data and Control Chart Analysis*, 6th edition, MNL 7 (ASTM International, West Conshohocken, PA, 1991)

2 ISO Central Secretariat, 1, rue de Varembé, Case postale 56, CH-1211 Genève 20, Switzerland, Website: http://www.iso.ch; ISO standards are available in the United States through the American National Standards Institute, Website: http://www.ansi.org, and in most other countries through the ISO member body.

3 International Vocabulary of Basic and General Terms in Metrology, Second Edition [VIM] (ISO, Genève, 1993)

SUPERSEDED

Documents
Document Sharing Hub

Full document details: SEMI E89-1106E

$\bar{x}$ = sample mean,
n = number of measurements, and
x_i = i^{th} measurement value.

5.3.3 *calibration* — set of operations that establish the relationship between values of quantities indicated by a measurement system (MS) and the corresponding values assigned to reference materials.

5.3.3.1 *Discussion* — The purpose of calibration is to reduce or eliminate bias in the MS.

5.3.4 *certified reference material (CRM)* — reference material, one or more of whose property values are certified by a technically valid procedure, accompanied by or traceable to a certificate or other documentation issued by a certifying body.

5.3.5 *coefficient of variation* (*CV*) — population standard deviation expressed as a percentage of the mean value.

5.3.5.1 *Discussion* — CV can be estimated from the sample standard deviation, s, and the sample mean, $\bar{x}$, of a distribution as follows:

$$CV = \frac{s}{\bar{x}} \times 100 \tag{2}$$

CV is an appropriate measure of variability only when the sample standard deviation is proportional to the mean; otherwise it varies with the value of the measurand. If the sample standard deviation is independent of the value of the measurand, it is more appropriate to use it directly rather than CV.

5.3.6 *effect* — change in the expected value of a given response due to the change of a given factor from one level to another. It is a measure of influence that a particular variable level has on the output variable.

5.3.6.1 *fixed effect* — variable for which estimates of the mean are obtained for each level.

5.3.6.2 *random effect* — variable for which estimates of the mean are not obtained for each level; rather the variable is treated as a variance component.

5.3.7 *factor* — predictor variable whose level is changed with the intent of assessing its effect on the response variable (in a designed experiment) [adapted from ISO 3534-3].

5.3.7.1 *crossed factor(s)* — two factors are crossed when every level of one factor appears with every level of the second factor.

5.3.7.2 *fixed factor* — factor that has either all of its levels represented in an experiment or levels selected by a nonrandom process.

5.3.7.3 *nested factor(s)* — factor that has a different set of levels appearing within each level of a second factor. Factor B is nested in factor A when randomization of the levels of factor B is restricted to specific levels of factor A.

5.3.7.4 *random factor* — factor that has randomly sampled levels from a population of levels.

5.3.8 *gage* — alternate spelling of gauge.

5.3.9 *gauge* — instrument used to assign a value to a quantitative or qualitative characteristic of a physical entity or phenomenon.

5.3.10 *interaction* — effect for which the apparent influence of one factor on the response variable depends upon one or more other factors [ISO 3534-3].

5.3.11 *level* — value of a factor (in a designed experiment) [adapted from ISO 3534-3]. Also called "setting of a variable."

5.3.12 *linearity* — absence of changes in variability or bias as measurements are made at different points within the measurement range.

5.3.12.1 *Discussion* — Traditional definitions of linearity ignore the fact that variability can change over the measurement range, as well as bias. The assumption of constant variability over the measurement range should be verified during the MS analysis.

5.3.13 *lower specification limit (LSL)* — value of an attribute below which a product is said to be nonconforming.

5.3.14 *matching tolerance* (Δ_m) — difference in bias for any two measurement systems (MSs) of the same kind made under the conditions of reproducibility.

5.3.15 *measurand* — particular attribute of a phenomenon, body or substance subject to measurement. [VIM]

SUPERSEDED

Documents Document Sharing Hub
Full document details: SEMI E89-1106E

5.3.16 *measurement resolution, of a gauge* — smallest difference in measurand that can be meaningfully distinguished by the gauge.

5.3.17 *measurement subsystem* — any set of entities, processes, or conditions that share a common purpose in the measurement.

5.3.17.1 *Discussion* — A measurement subsystem may contain one or more of its own subsystems. For example, a wafer handling mechanism may be further composed of wafer loading and wafer positioning subsystems.

5.3.18 *measurement system* (*MS*) — all entities, procedures, and conditions that can influence the test result obtained with a given measurement process.

5.3.18.1 *Discussion* — The MS may include, but is not limited to, the gauge, operators, setup mechanics, wafers, locations on a wafer, environmental conditions, software used by the gauge, measurement method, etc. The MS may be comprised of measurement subsystems.

5.3.19 *measurement system analysis (MSA)* — procedure in which relevant sources of bias and variability associated with a measurement system (MS) are estimated.

NOTE 2: MSA is also sometimes called gauge (or gage) repeatability and reproducibility (GRR or GR&R).

5.3.20 *nested design* — experimental design in which different levels of one factor appear in each level of a second factor.

5.3.21 *population standard deviation* (σ) — square root of the population variance.

5.3.22 *population variance* (σ^2) — measure of dispersion associated with a population distribution.

5.3.22.1 *Discussion* — For continuous distributions, the population variance is the second central moment.

5.3.23 *precision* — general estimator of the variability of a measurement process about the mean value of the test results obtained.

5.3.23.1 *Discussion* — Precision is a random component of measurement uncertainty. Unless the measurement process is in a state of statistical control, the precision of the process has no meaning. Since the precision is poorer for greater dispersion of the test results, specific measures of variability (such as repeatability and reproducibility) are actually direct measures of the imprecision of the measurement process.

5.3.24 *precision-to-tolerance* (*P/T*) *ratio* — ratio of the precision of a measurement system (MS) to the tolerance (i.e., absolute magnitude of the full range of the product specification).

5.3.24.1 *Discussion* — If the variability associated with the measurement of a parameter by an MS is very small compared with the width of the specification range, the probability of obtaining a test result outside the specification limits when the value of the parameter actually lies within the specification limits (or conversely) is quite small. On the other hand, if the ratio is too large, the probability of obtaining a false test result is much greater.

5.3.25 *predictor variable* — variable that can contribute to the explanation of the outcome of a designed experiment. Also called "input variable," "descriptor variable," and "explanatory variable."

5.3.25.1 *Discussion* — The term "independent variable" is not recommended as a synonym due to potential confusion with independence.

5.3.26 *product standard deviation* ($\sigma_{Product}$) — population standard deviation associated with the distribution of values of all possible realizations of a property of an entity manufactured under specified conditions.

5.3.26.1 *Discussion* — The product variability may be estimated by taking a representative sample from the population and calculating its sample standard deviation ($s_{Product}$) taking suitable account of MS variation (see ¶ 10.2).

5.3.27 *reference material* — material or substance, one or more of whose property values are sufficiently homogeneous and well established to be used for the calibration of a MS, for the assessment of a measurement method or for assigning values to materials.

5.3.28 *repeatability* (σ_r) — variability associated with repeated measurements taken under repeatability conditions.

5.3.29 *repeatability conditions* — test conditions involving acquisition of a series of test results with the same test protocol and MS setup in the same laboratory by the same operator on the same equipment in the shortest practical period of time on the same test wafer without explicit recalibration.

5.3.29.1 *Discussion* — The acquisition of test data under repeatability conditions is intended to avoid influences of long-term drift, operator or MS differences, material variability, and the like. Recalibration of the MS is expected to

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E89-1106E

cause discontinuous differences in test results. However, if recalibration is required by the test protocol or is internal to the MS, it is considered to be an allowable variation in determination of repeatability.

5.3.30 *reproducibility (σ_R)* — variability associated with the measurement system (MS) when measurements are made under different (but typical) conditions.

5.3.30.1 *Discussion* — Changes associated with subsystems or test conditions are potential sources of variation to be estimated. Repeatability is one source of variation. Other relevant sources of variability may include time, operator, setup procedure, wafer (of like variety), measurement location, test instrumentation, environmental conditions, etc. Although the total number of contributors to the variance can be exceedingly large, one typically focuses on a subset that accounts for a significant portion of the expected MS variability. For clarity, the selected subset should be reported together with the reproducibility. If q different conditions introduce variability into the measurement independently from one another, the variances add directly

$$\sigma_R = \sqrt{\sigma_1^2 + \sigma_2^2 + \sigma_3^2 + \ldots + \sigma_q^2} \tag{3}$$

and they may be separated by the use of judiciously designed experiments.

5.3.31 *response variable* — variable representing the outcome of a designed experiment. Also called "output variable."

5.3.31.1 *Discussion* — The term "dependent variable" is not recommended as a synonym due to potential confusion with independence.

5.3.32 *root sum of squares (RSS) difference* — square root of the difference of the squares of two numbers.

5.3.33 *root sum of squares (RSS) sum* — square root of the sums of the squares of two or more numbers.

5.3.34 *sample standard deviation (s)* — square root of the sample variance.

5.3.35 *sample variance (s^2)* — measure of dispersion given by the average squared deviation from the mean for a set of numbers.

5.3.35.1 *Discussion* — If x_i is an individual measurement, $\bar{x}$ is the average across all measurements, and n is the number of measurements, then

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2 \tag{4}$$

The denominator value of $n - 1$ is used instead of n to make the sample variance an unbiased estimator of the population variance.

5.3.36 *signal–to-noise ratio (SNR)* — ratio of the variation in the manufactured product to the precision of the measurement system (MS).

5.3.36.1 *Discussion* — Because it is difficult to directly measure the standard deviation of the product without including variation due to the measurement instrument, SNR is generally defined as:

$$SNR = \sqrt{\frac{\sigma_{Total}^2 - \sigma_R^2}{\sigma_R^2}} \tag{5}$$

where σ_{Total}^2 is an estimate of the total population variance obtained from appropriate measurements of a large, representative sample of the product.

5.3.37 *stability* — absence of additional variability due to taking measurements over time (typically several days or longer).

5.3.38 *statistical model* — mathematical function relating one or more variables to known and measurable inputs plus one or more unknown stochastic (error) terms.

5.3.38.1 *Discussion* — A statistical model consists of three parts. The first part is the response variable that is being modeled. The second part is the deterministic or the systematic part of the model that includes predictor variables. Finally, the third part is the random error or stochastic part of the model, which can be quite elaborate. An example of a statistical model is:

SUPERSEDED

Documents Document Sharing Hub
Full document details: SEMI E89-1106E

$$y_j = \sum_{i=1}^{p} \beta_i x_i + e_j \tag{6}$$

where:

y_j = j^{th} measurement,
p = number of input variables,
x_i = i^{th} input variable,
β_i = its corresponding coefficient, and
e_j = error associated with the jth measurement.

In many cases the error distribution(s) are specified before the model is fit (e.g., as normal).

5.3.39 *tolerance* — absolute magnitude of the full range of the product specification.

5.3.40 *total variance* (σ^2_{Total}) — sum of the product variance and the square of the reproducibility.

5.3.41 *uncertainty* — parameter, associated with a measurement, that characterizes the dispersion of values that can be reasonably attributed to the object being measured.

5.3.41.1 *Discussion* — Two types of measurement uncertainty are:

- Type A: uncertainty components evaluated by statistical methods and
- Type B: uncertainty components evaluated by other than statistical methods.

5.3.42 *upper specification limit (USL)* — value of an attribute above which a product is said to be nonconforming.

5.3.43 *variable* — quantitative or qualitative characteristic of an object, processes, or state that may take on more than one value.

5.3.43.1 *Discussion* — When the values occur unpredictably, it is a random variable.

5.3.44 *variance* — population variance (see ¶ 5.3.22).

6 Goals of and Preliminary Steps in Planning an MSA

6.1 One major goal of an MSA is to determine MS reproducibility under the desired operating conditions. This may range from a simple repeatability study, to a complex determination of the relative contributions of many sources of MS variability using factorial experiments and analysis of variance (ANOVA). Some of the typical types of variability-determination goals to be met by an MSA include the following:

6.1.1 Determination of repeatability.

6.1.2 Determination of the effect of loading and unloading the wafer samples between measurements.

6.1.3 Determination of stability.

NOTE 3: Stability may also have a bias component that can be confounded with the variability component of stability (see ¶ 6.2.3).

6.1.4 Assessment of the multiple factors that affect the MS variability, which requires the use of factorial experiments and ANOVA.

6.1.5 Assessment of the largest sources of variability in order to focus improvement efforts on the most critical sources. This identification can be made from the analysis of a properly designed and executed factorial experiment.

6.1.6 Once the reproducibility is determined over the desired conditions, secondary metrics, such as P/T ratio and SNR can be easily calculated.

6.2 A second major goal of an MSA is to determine bias and bias-related considerations. In general, bias studies require standards or "golden wafers" with "known" true measurement values or agreed-upon consensus values. Corrective actions ranging from simple recalibration to repair of the MS may be required as a result of the bias determinations. Some of the typical elements of a bias study include:

6.2.1 Determination of bias.

6.2.2 Determination of linearity or whether bias is constant over the range of values that the metrology equipment will be used.

NOTE 4: Linearity also may have a variability component.

6.2.3 Determination of the stability or whether the bias is constant over time.

NOTE 5: Stability also may have a variability component (see ¶ 6.1.3).

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E89-1106E

6.2.4 Determination of matching tolerance.

NOTE 6: In the absence of certified or other reference materials, as required for a determination of bias, matching tolerance may be determined by measuring the same set of test materials on the two MSs whose matching is being evaluated. Although this alternate method does not provide a measure of the value of bias, the difference between the mean values measured on the two MSs is the same as the difference in bias.

6.3 Before undertaking an MSA, determine and list the specific goals to be attained by the MSA to be performed.

6.4 It must be emphasized that all of the procedures in this guide yield meaningful estimates only on instruments that are in statistical control. Therefore the first step in analyzing the capability of any MS is to establish that it is in statistical control through control charts or other accepted quality control techniques. Detailed discussion of these techniques is outside the scope of this guide (see ¶ 3.1).

6.5 It is also useful to review prior studies on the MS to be analyzed to determine (1) what sources of variability were significant and (2) previously obtained values for quantities such as bias, repeatability, and reproducibility.

7 Procedure to Determine Reproducibility

7.1 Identify all the factors of interest that contribute to non-negligible sources of variability in the MS. These factors should be made explicit before an experiment can be appropriately designed or accurately analyzed. The most common sources include:

- Repeatability,
- Load-unload operation,
- Time over intervals of day(s) or week(s),
- Wafer handling hardware, including ability to place the measurement probe at the same location on the wafer,
- Measurement software, including type, revision, or both,
- Setup procedure,
- Environmental conditions, including temperature and humidity,
- Shift (may be confounded with environmental conditions or time or both),
- Magnitudes of attribute(s) of test wafer(s) under investigation, and
- Operator (usually not significant for automated MSs).

NOTE 7: The factors listed are all included under the term "reproducibility." There is no common agreement as to what factors should be included in reproducibility. Therefore the selection and recording of the factors included in a particular MSA are especially critical.

NOTE 8: Variability associated with factors independent of the MS, such as true value differences in the item being measured, should not be included in reproducibility. Given the design of the MS, it may be necessary to estimate these sources of variability before they are removed from total variability. If such factors are included in the MS, as for example, when different sites on a wafer are measured, variability associated with these factors should be estimated and removed from total MS variability. Although variability due to the measurement process physically altering the measurand (i.e., the interaction between MS and wafer) should be included in reproducibility, consideration of this topic is beyond the scope of this guide (see ¶ 3.3).

7.2 Identify all factors outside of the MS for which an interaction with one or more factors within the MS may exist. Treat these factors as fixed factors and the interactions as random factors.

NOTE 9: Typically, the only fixed factor considered is related to the measured object, where making measurements at different locations, for different true values, or on different wafer types may be influenced by certain aspects of the MS.

7.3 Identify the nested factors and the factors in which they are nested. Several levels of nesting may exist for a given factor.

NOTE 10: Nesting occurs when the measurement procedure causes a natural hierarchy of events to exist. For example, if repeatability is estimated by taking measurements with the wafer fixed in the gauge, and variability due to the load-unload operation is estimated by taking a second set of measurements after the wafer has been removed and reloaded, repeatability is nested in load-unload. The nesting occurs because each set of repeated measurements is unique to the load in which they appear. Nesting may also occur if a physical hierarchy exists, such as when both measurement locations (sites) and wafer type are fixed factors in an MSA (sites being nested in wafer type).

7.4 Select appropriate test wafers for the MSA.

7.4.1 In the most general case, many types of wafers should be used. These should be representative of the types and levels to be measured in practice by the MS under evaluation.

SUPERSEDED

Documents Document Sharing Hub
Full document details: SEMI E89-1106E

7.4.2 The test wafers should be as uniform as possible to prevent confounding of handling, loading, or positioning factors, if these are chosen to be examined.

7.4.3 It is necessary to use only one test wafer of each type selected.

7.4.4 Use the same test wafers throughout the MSA.

7.5 Prepare a statistical model for the MSA in order to plan and analyze it properly.

NOTE 11: Statistical models are discussed in Related Information 3. Eastman[4] gives some examples of models for reproducibility studies.

7.6 Select the measurement protocol. Include several examples of every factor that was selected in ¶ 7.1.

7.7 Select and implement the appropriate MS setup.

NOTE 12: Usually, only a single setup is used in an MSA to determine reproducibility. However, if setup procedure has been chosen as a factor to be evaluated, select the various setups to be used in the MSA.

NOTE 13: Examples of MSAs are given in Related Information 4, 5, and 6. Additional examples may be found in the literature.[4]

7.8 Determine if the MS needs to be calibrated.

7.8.1 Calibrate, if necessary, following the manufacturer's recommended procedure.

7.8.2 Do not recalibrate during the MSA unless required by the normal operating procedure.

7.9 Perform the MSA, randomizing both the sample wafers chosen and the order in which measurements are made, whenever possible. Record the actual order in which all measurements are made.

7.10 To correctly estimate the reproducibility and individual sources of variance, conduct a variance components analysis.

7.10.1 If the experiment has been designed to include multiple repeated measurements taken under the same (repeatability) conditions, it is possible to estimate repeatability. Without multiple repeated measurements, repeatability will be confounded with other sources of variability.

7.10.2 Other factors are those from ¶ 7.1 that have been included in the experiment design.

7.10.3 Reproducibility is the square root of the sum of the individual variance components (see Equation 3 in ¶ 5.3.30.1).

NOTE 14: The exact nature of the variance components analysis depends on the design of the experiment conducted. For details, see the examples in Related Information 4, 5, and 6; Box, Hunter, and Hunter;[5] or Montgomery.[6] In designing and performing a complete designed experiment and the associated analysis, it is strongly recommended to utilize some form of commercially available statistical analysis software. If such a product is used, it is advisable to confirm that the assumptions made in the software are congruent with those in this guide.

8 Procedure to Determine Bias, Including Linearity, Stability, and Matching Tolerance

8.1 Obtain *J* appropriate reference materials, preferably CRMs, where *J* is between 3 and 10 so that the entire range of interest of the parameter is covered. If CRMs are not available, use wafers with well accepted consensus values covering the range of values of the attributes to be measured. The reference materials should have characteristics similar to the wafers to be measured by the MS being evaluated or their values should be transferable to such wafers.

NOTE 15: Bias cannot be determined if suitable reference materials are not available. In such situations, it is necessary to resort to correlation experiments to establish the systematic errors between different MSs. If the bias is not known, these systematic errors may confound the estimated reproducibility.

8.2 Select and implement the appropriate MS setup.

8.3 Determine if the MS needs to be calibrated. Calibrate, if necessary, following the manufacturer's recommended procedure. Do not recalibrate during the bias determination unless required by the normal operating procedure.

4 Eastman, S. A., "Evaluating Automated Wafer Measurement Instruments," International SEMATECH technology transfer document 94112638A-XFR (February 28, 1995), Section 5 and case studies in appendices. However, note that neither the terminology nor the nomenclature in this document is identical with that given in this guide. A PDF file of this report can be downloaded from International SEMATECH's public website at http://ismi.sematech.org/docubase/abstracts/2638axfr.htm.

5 Box, G. E. P., Hunter, W. G., and Hunter, J. S., Statistics for Experimenters (Wiley, New York, 1978)

6 Montgomery, D.C., Design and Analysis of Experiments, 5th Ed. (Wiley, New York, 2000)

SUPERSEDED
Documents
Document Sharing Hub
Full document details: SEMI E89-1106E

8.4 Chose the minimum number of measurements to be made on each of the J reference materials, n, as the smallest integer greater than $(4\sigma/\delta)^2$ where σ is the estimate of reproducibility (σ_R, see § 7) and δ is the bias shift to be detected.

NOTE 16: In this relationship, the assumption is made that the maximum acceptable probability of calibrating when unnecessary (Type 1 error) is 0.1 and the maximum tolerable probability of not calibrating when necessary (Type 2 error) is 0.01. If use of other values of these errors is desired, the more complete formula for minimum sample size is the smallest integer greater than $[\sigma(z_1 + z_2) / \delta]^2$ where:

z_1 = value above which $p_1/2$ percent of the standard normal distribution falls where p_1 is the largest acceptable risk of calibrating when unnecessary,
z_2 = value above which $p_2/2$ percent of the standard normal distribution falls where p_2 is the largest acceptable risk of not calibrating when necessary,
σ = estimate of the variability as defined in ¶ 8.4, and
δ = shift (bias) to be detected.

8.4.1 If no information on MS variability is available, use a minimum of 16 measurements on a reference wafer to estimate bias.

8.5 Perform the measurements on one or more days, depending if stability information is desired or not. To minimize systematic errors, randomize the sequence of measurements among the reference wafers for each daily measurement. If the MS is performed over several days, make an approximately equal number of measurements on each day with at least two measurements made per day.

8.6 *Analysis of the Measurement Data*

8.6.1 This analysis can be done with most commercially available statistical analysis software packages. The procedures given in ¶ 8.7 are intended only for those without access to such software.

8.6.2 The full model to determine the effect of bias over time and measurement range is:

$$E[y_{ij}] = b_i + a_i x_{ij} \tag{7}$$

where
$E[y_{ij}]$ = expected response for reference wafer j on day i,
b_i = bias on day i,
x_{ij} = certified value for reference wafer j on day i, and
a_i = slope for day i.

NOTE 17: If multiple measurements are made for a given reference wafer on a given day, it is assumed that

$$E[y_{ijk}] = E[y_{ij}] \tag{8}$$

where k is any one of the multiple measurements made on reference wafer j on day i.

8.6.3 When bias and slope do not differ significantly over days (i.e., $a_i \approx a$ and $b_i \approx b$ for all i), the reduced model is:

$$E[y_j] = b + ax_j \tag{9}$$

8.6.4 *Linearity* — The calibration coefficient a is a linear constant that relates the rate at which bias changes for every unit change in the measurement range. This value should not be significantly different from one. Otherwise, the MS exhibits nonlinearity.

8.6.5 *Stability* — Test for a significant difference between the full and reduced models given by Equations 7 and 9, respectively. If there is no significant difference, continue. Otherwise, bias is not stable over time.

8.6.6 Test for model lack of fit. Significant lack of fit indicates a possible nonlinear relationship involving bias that is not captured by the reduced model.

8.6.7 If the test for lack of fit is not significant, estimate a and b.

8.6.7.1 If a is not statistically different from one and b is statistically different from zero, b is an estimate of the bias.

8.6.7.2 If a is statistically different from one, bias is not constant. A linear relationship exists between bias and measurement range.

8.6.7.3 If b is not statistically different from zero, the estimate of bias is zero.

8.7 *Detailed Analysis Procedures* — If the measurements are made over a single day only, skip ¶¶ 8.7.1 and 8.7.2 and begin the analysis with ¶ 8.7.3. Also skip ¶ 8.7.5. In this case, no information about time variation of bias (stability) can be obtained from the MSA. Otherwise begin the analysis with ¶ 8.7.1.

SUPERSEDED

Documents Document Sharing Hub
Full document details: SEMI E89-1106E

8.7.1 Estimate the slopes ($\tilde{a}_i$) and biases ($\tilde{b}_i$) for each day's measurements to determine the effect of time and measurement range on the bias:

$$\tilde{a}_i = \frac{\sum_{j=1}^{J} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y}_i)(x_{ijk} - \bar{x}_i)}{\sum_{j=1}^{J} \sum_{k=1}^{n_{ij}} (x_{ijk} - \bar{x}_i)^2} \tag{10}$$

$$\tilde{b}_i = \bar{y}_i - \tilde{a}_i \bar{x}_i \tag{11}$$

where:

J = number of reference wafers,
n_{ij} = number of measurement results for reference wafer j on day i,
y_{ijk} = result of measurement k on reference wafer j on day i,
$\bar{y}_i$ = mean of the measurement results on day i,
x_{ijk} = value of the reference parameter on day i, and
$\bar{x}_i$ = mean of the values of the reference parameter on day i.

8.7.2 Calculate the expected response on each reference wafer j on each day i, $E[y_{ij}]$:

$$E[y_{ij}] = \tilde{b}_i + \tilde{a}_i x_{ij} \tag{12}$$

where:

x_{ij} = value of the reference parameter for wafer j of day i, and
$\tilde{a}_i$ and $\tilde{b}_i$ are the slope and intercept (bias) for day i as found from Equations 10 and 11, respectively, in ¶ 8.7.1.

NOTE 18: When multiple measurements are made for a given reference wafer on a given day, it is assumed for all k that

$$E[y_{ijk}] = E[y_{ij}] \tag{13}$$

where k is any one of the multiple measurements made on the i day on the j reference wafer.

8.7.3 Estimate the slopes ($\hat{a}_i$) and biases ($\hat{b}_i$) for the entire data set:

$$\hat{a} = \frac{\sum_{i=1}^{D} \sum_{j=1}^{J} \sum_{k=1}^{n_{ij}} (y_{ijk} - \bar{y})(x_{ijk} - \bar{x})}{\sum_{i=1}^{D} \sum_{j=1}^{J} \sum_{k=1}^{n_{ij}} (x_{ijk} - \bar{x})^2} \tag{14}$$

$$\hat{b} = \bar{y} - \hat{a}\,\bar{x} \tag{15}$$

where:

D = number of days over which measurements are made,
$\bar{x}$ = grand mean of the values of the reference parameter,
$\bar{y}$ = grand mean of the measurement results, and

the other parameters are defined in ¶ 8.7.1.

NOTE 19: This procedure makes the assumption that the bias and slope do not differ significantly over days. This assumption is tested in ¶ 8.7.5.

8.7.4 Calculate the expected response on each wafer j, $E[y_j]$:

$$E[y_j] = \hat{b} + \hat{a}\,x_j \tag{16}$$

8.7.5 Determine whether bias and slope differ significantly over days or not as follows:

8.7.5.1 Calculate the Sum of Squares for the difference between the two cases, SS_M:

$$SS_M = \sum_{i=1}^{D} \sum_{j=1}^{J} \left(E[y_{ij}] - E[y_j]\right)^2 \tag{17}$$

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E89-1106E

8.7.5.2 Calculate the Error Sum of Squares (SS_e):

$$SS_e = SS_Y - SS_\gamma - SS_M \tag{18}$$

where:

$$SS_Y = \sum_{i=1}^{D}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}}\left(y_{ijk} - \bar{y}\right)^2 \tag{19}$$

and

$$SS_\gamma = \sum_{i=1}^{D}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}}\left(y_{ijk} - \bar{y}\right)\left(x_{ijk} - \bar{x}\right) \tag{20}$$

8.7.5.3 Calculate probability associated with the function F_M that follows a Fisher's F distribution with $2D - 1$ and $n - 2D$ degrees of freedom under the hypothesis that there is no difference between the two cases:

$$F_M = \frac{(n-2D)SS_M}{(2D-1)SS_e} \tag{21}$$

NOTE 20: This probability can be calculated using the Excel function, FDIST(F_M,$2D$–1, n–$2D$).

8.7.5.4 Reject the hypothesis that there is no difference between the two cases when the probability associated with F_M is small (e.g., 0.05 or below). In this case, bias is not stable over time and it is not possible to establish a value for it.

8.7.6 Test for lack of fit as follows:

8.7.6.1 Calculate the Sum of Squares for Pure Error (SS_p):

$$SS_p = \sum_{i=1}^{D}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}}(y_{ijk} - \bar{y}_j)^2 \tag{22}$$

where all the symbols have been defined previously.

8.7.6.2 Calculate the lack of fit F test

$$F_{LOF} = \frac{(J-2)SS_p}{(n-J)(SS_e - SS_p)} \tag{23}$$

where, again, all the symbols have been defined previously.

8.7.6.3 Under the assumption of no lack of fit, the ratio F_{LOF} follows Fisher's F distribution. Calculate the probability associated with this ratio.

NOTE 21: This probability can be calculated using the Excel function: FDIST(F_{LOF}, n–J, J–2).

8.7.6.4 Reject the assumption of no lack of fit if this probability is small (e.g., 0.01 or less). In this case, nonlinearity is probably present in the MS.

8.7.7 If there is no evidence of lack of fit, calculate the approximate 95% confidence interval for $\hat{a}$:

$$\hat{a} \pm 2\sqrt{\frac{SS_e/(n-J)}{\sum_{i=1}^{D}\sum_{j=1}^{J}\sum_{k=1}^{n_{ij}}(x_{ijk} - \bar{x}_j)^2}} \tag{24}$$

where:

$\bar{x}_j$ = (mean) value of the reference parameter from reference wafer j.

8.7.7.1 If the interval contains the number one, $\hat{a}$ is not statistically different from one. If this is the case and also $\hat{b}$ is significantly different from zero, $\hat{b}$ is the estimate of the bias.

8.7.7.2 If the interval does not contain the number one, $\hat{a}$ is significantly different from one and bias is not constant with parameter level.

8.7.7.3 If $\hat{b}$ is not significantly different from zero, the estimate of bias is zero.

SUPERSEDED

Documents Document Sharing Hub
Full document details: SEMI E89-1106E

8.8 *Matching Tolerance*

8.8.1 Determine the bias of two MSs of the same kind under conditions of reproducibility.

8.8.2 If the results give a stable bias determination for each MS and if each MS has acceptable linearity, subtract the two biases to obtain the matching tolerance, Δ_m:

$$\Delta_m = bias_1 - bias_2 \tag{25}$$

where:

$bias_1$ = bias of the first MS and
$bias_2$ = bias of the second MS.

9 Procedure to Determine P/T Ratio

9.1 Estimate the precision of the MS, P, as $6s_R$ if the specification is symmetrically two-sided or as $3s_R$ if the specification is asymmetrical or one-sided.

9.2 Take the tolerance, T, as follows:

9.2.1 The difference between the USL and the LSL for a symmetrical two-sided specification,

9.2.2 The smaller difference between the USL (or LSL) and the target value for an asymmetrical two-sided specification, or

9.2.3 The difference between the USL (or LSL) and the median of the expected distribution for a one-sided specification.

NOTE 22: If a symmetrical two-sided specification is given as

Target ± Tolerance,

double the stated tolerance to obtain the correct value.

9.3 Calculate the P/T ratio, in percent, as follows:

$$P/T(\%) = \frac{P}{T} \times 100 \tag{26}$$

rounding to the nearest percent.

NOTE 23: Generally, a P/T ratio should be 30% or less.

10 Procedure to Determine *SNR*

10.1 Estimate the precision of the MS, P, as $6s_R$ if the process distribution is two-sided or as $3s_R$ if the process distribution is asymmetrical or one-sided.

10.2 Estimate the population standard deviation of the process distribution from process data as $\sigma_{process}$. Because it is difficult to directly measure the standard deviation of the product without including variation due to the MS, $\sigma_{process}$ is generally defined as:

$$\sigma_{process} = \sqrt{\sigma_{Total}^2 - \sigma_R^2} \tag{27}$$

where:

σ_{Total}^2 = variance obtained by measuring a large representative sample of the product and

σ_R^2 = variance determined from the reproducibility (see § 7).

NOTE 24: Detailed procedures for obtaining the process distribution and the associated standard deviation are outside the scope of this guide.

10.3 Calculate the *SNR* as follows:

$$SNR = \frac{\sigma_{process}}{\sigma_R} \tag{28}$$

rounding to the nearest percent.

NOTE 25: In general, an *SNR* of 10 or more generally means the MS is suitable for measuring the product, while an *SNR* of less than 3 or 4 may be a concern in a particular measurement process.

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E89-1106E

11 Related Documents

11.1 The following documents describe MSA methodologies similar to those advocated in this guide:

Ballard, D. H., McCormack, D. W., Jr., Moore, T. L., Pore, M., Prins, J., Tobias, P. A. (1998). "A Comparison of Gauge Study Practices," Proceedings from the 1997 Joint Statistical Meetings of the American Statistical Association, Quality and Productivity Section

John, Peter, "Alternative Models for Gauge Studies," International SEMATECH technology transfer document 93081755A-TR (February 24, 1994); PDF file can be downloaded from International SEMATECH's public website at http://www.sematech.org/docubase/wrappers/26.htm

Montgomery, D.C. and Runger G. C. (1993), "Gauge Capability and Designed Experiments. Part I: Basic Methods," *Quality Engineering* **6**(1), 115–135

Montgomery, D.C. and Runger G. C. (1993), "Gauge Capability and Designed Experiments. Part II: Experimental Design Models and Variance Component Estimation," *Quality Engineering* **6**(2), 289–305

Potter, R.W. (1991), "Measurement System Capability Analysis," IEEE/SEMI Advanced Semiconductor Manufacturing Conference, pp. 121–125

---, Measurement Systems Analysis Reference Manual, Third Edition, (2002); Daimler Chrysler Corp., Ford Motor Co, and General Motors Corp. Automotive Industry Action Group (AIAG); http://www.aiag.org

---, "NIST/SEMATECH e-Handbook of Statistical Methods": http://ismi.sematech.org/docubase/abstracts/1755atr.htm

SUPERSEDED

Documents Document Sharing Hub

Full document details: SEMI E89-1106E

RELATED INFORMATION 1
MEASUREMENT UNCERTAINTY CALCULATIONS

NOTICE: This related information is not an official part of SEMI E89. It was derived from task force deliberations during the revision of SEMI E89-0999 in 2001-2003. This related information was approved by full letter ballot procedures and was approved for publication by the NA RSC on August 16, 2004.

R1-1.1 Measurement uncertainty depends on the repeatability of the MS, the reproducibility of the result over time, the number of measurements in the test result, and all sources of random and systematic error that could contribute to disagreement between the measurement result and its reference value.

R1-1.2 Determine the measurement uncertainty according to the following rules:

- *Expressing Uncertainty* — Each uncertainty component is quantified by a standard deviation.
- *Bias* — All biases are assumed to be corrected and any uncertainty due to bias is the uncertainty of the correction.
- *Standard Uncertainty* — All uncertainty components (standard deviations), whether Type A or Type B (see § 5.3.41), are combined as an RSS sum to arrive at a "standard uncertainty," u. This standard uncertainty is the standard deviation of the reported value, taking into account all sources of error, both random and systematic, that affect the measurement result.
- *Extended Uncertainty* — If the purpose of the uncertainty statement is to provide coverage with a high level of confidence, an expanded uncertainty is computed as $U = ku$, where $k = 2$ is typically chosen for an approximate 95% coverage.
- *Uncertainty Interval* — If Y is the reported measurement value, then the symmetric interval from $Y - U$ to $Y + U$ is the uncertainty interval associated with the measurement value.

R1-1.3 The measurement result for which an uncertainty is required must be completely specified as to

- the number of repetitions that were averaged,
- the test method,
- the environmental conditions,
- the operating conditions over which the repetitions were made, and
- any calibration uncertainty.

R1-1.4 Obtain Type A uncertainty for a particular measurement from a determination of reproducibility (see § 7) conducted according to the same specifications as the particular measurement and including all relevant components of reproducibility.

SUPERSEDED

Documents
Document Sharing Hub
Full document details: SEMI E89-1106E

RELATED INFORMATION 2
TESTING MEASUREMENT DISTRIBUTIONS FOR NORMALITY AND EQUAL REPEATABILTY

NOTICE: This related information is not an official part of SEMI E89. It was derived from task force deliberations during the revision of SEMI E89-0999 in 2001-2003. This related information was approved by full letter ballot procedures and was approved for publication by the NA RSC on August 16, 2004.

R2-1 Testing for Normality

R2-1.1 Testing for normality can be achieved by a commonly applied statistical hypothesis test. In this test, the condition under consideration is that the sample being examined comes from the normal distribution. While there are several ways of implementing this procedure, in each case assume that the measurement results come from a normal distribution and then calculate the probability of obtaining the particular observed sample. Before calculating this probability value (*p*-value), an *a priori* decision is made as to how small this *p*-value has to be to warrant the conclusion that the hypothesis is unreasonable and hence it is unlikely that the observed distribution is normal. A level of 0.05 is commonly chosen and the sample is then tested. If the *p*-value is greater than 0.05 (or the chosen test level, if different from 0.05), one accepts that the underlying distribution is normal.

R2-1.2 Statistical tests for normality can be found in most commercial statistical software. Four common tests are Andersen-Darling, Pearson's Chi Square, Kolmogorov-Smirnov, and Shapiro-Wilk.

R2-1.3 As an example, the 25 observations shown in Table R2-1 are a random sample taken from a manufacturing process. A commonly used statistical package was used to generate the statistical tests mentioned above and the results obtained are shown in Table R2-2.

Table R2-1 Measurement Data

481	476	471	477	479
483	477	480	475	479
480	473	479	484	473
481	475	478	472	480
485	480	478	475	481

R2-1.4 Note that if we require a *p*-value is less than or equal to 0.05 as a rejection criterion, none of these statistical procedures would reject normality for the manufacturing process measurements although the calculated *p*-values are all different.

R2-1.5 For some of these tests, software packages may rely on a combination of tabled values and interpolation to arrive at the answer. For that reason, the numbers given for the same test may vary depending on the software used.

Table R2-2 Calculated Normality Statistics

Test	*Test Statistic Value*	*Calculated p-Value*
Anderson-Darling	0.28726858	>0.250
Chi-Square	2.86169631	0.239
Kolmogorov-Smirnov	0.11888191	>0.150
Shapiro-Wilk	0.973356	0.731

R2-2 Testing Two Measurement Systems for Equal Repeatability

R2-2.1 When comparing two MSs, it is possible to test if the repeatabilities are equal. To do so, calculate the ratio of the repeatability of each MS. When forming this ratio, place the larger value in the numerator. The hypothesis under consideration is that the repeatabilities are equal. When this is true, the ratio follows an *F* distribution with ν_1 degrees of freedom for the numerator and ν_2 degrees of freedom for the denominator.

SUPERSEDED

Documents Document Sharing Hub
Full document details: SEMI E89-1106E