Designation: F 1806 – 97 (Reapproved 2001)$^{\epsilon 1}$

Standard Practice for Tire Testing Operations–Basic Concepts and Terminology for Reference Tire Use[1]

This standard is issued under the fixed designation F 1806; the number immediately following the designation indicates the year of original adoption or, in the case of revision, the year of last revision. A number in parentheses indicates the year of last reapproval. A superscript epsilon (ϵ) indicates an editorial change since the last revision or reapproval.

ϵ^1 NOTE—Editorial corrections were made throughout the document in December 2001.

1. Scope

1.1 This practice presents some basic concepts for tire testing and a standard set of terms relating to the use of reference tires frequently used for comprehensive tire testing programs. The tests may be conducted in a laboratory on various dynamometer wheels or other apparatus as well as at outdoor proving ground facilities. The overall objective of this practice is to develop some elementary principles for such testing and standardize the terms used in these operations. This will improve communication among those conducting these tests as well as those using the results of such testing.

1.2 In addition to the basic concepts and terminology, a statistical model for tire testing operations is also presented in Annex A1. This serves as a mathematical and conceptual foundation for the terms and other testing concepts; it will improve understanding. The annex can also serve for future consultation as this practice is expanded to address additional aspects of the testing process.

1.3 This overall topic requires a comprehensive treatment with a sequential or hierarchical development of terms with substantial background discussion. This cannot be accommodated in Terminology F 538.

1.4 *This standard does not purport to address all of the safety concerns, if any, associated with its use. It is the responsibility of the user of this standard to establish appropriate safety and health practices and determine the applicability of regulatory limitations prior to use.*

2. Referenced Documents

2.1 *ASTM Standards:*

E 1136 Specification for a Radial Standard Reference Test Tire[2]

F 538 Terminology Relating to the Characteristics and Performance of Tires[3]

F 1082 Practice for Tires–Determining Precision for Test Method Standards[3]

F 1650 Practice for Evaluating Tire Traction Performance Data Under Varying Test Conditions[3]

3. Significance and Use

3.1 Tire testing operations usually consist of a sequence of tests that involve special "reference" tires in addition to the candidate tires being evaluated for their performance characteristics. Reference tires serve as an "internal benchmark" which may be used to adjust for variation in test results to give improved comparisons among the candidate tires. Numerous approaches have been adopted using different terminology for such testing. This causes confusion and the purpose of this practice is to standardize some of the elementary concepts and terminology on this topic.

4. Summary of the Practice

4.1 Elementary testing concepts, terms, and definitions are developed in hierarchical or sequential order beginning with basic testing operations. Each definition may be accompanied by a specific discussion or expanded text section appropriate to general definitions. Many of the terms could be defined as adjectives; however, as recommended by ASTM policy, the word "tire" is included in each definition avoiding the complication of defining adjectives. The definitions apply equally to items or objects other than tires.

5. Basic Testing Concepts and Terms

5.1 *Background on Testing*:

5.1.1 Despite the adoption of standardized testing procedures, test result variation influences data generated in any type of testing. As outlined in Annex A1, there are two main categories: [1] variation inherent in the production process for a group of nominally identical objects or tires and [2] variation due to the measurement operation. Each of these two sources may be further divided into two types of variation; [1] systematic or bias variation (the variation causing one laboratory to be consistently different from another laboratory) and [2] random error variation. Both types can exist simultaneously for either of the main categories.

5.1.2 Random variation can be reduced to a low level by appropriate replication and sampling procedures, but bias

[1] This practice is under the jurisdiction of Committee F09 on Tires and is the direct responsibility of Subcommittee F09.94 on Terminology.

Current edition approved May 10, 1997. Published April 1998.

[2] *Annual Book of ASTM Standards*, Vol 04.03.

[3] *Annual Book of ASTM Standards*, Vol 09.02.

variation cannot be so reduced. Bias variation can be reduced or eliminated by the appropriate use of reference objects or tires. This is the major rationale for their use in testing operations (see Annex A1).

5.1.3 Bias variation can also be reduced or eliminated by comprehensive programs to sort out causes of such perturbations and eliminate these causes.

5.2 *Elementary Testing Terms*:

5.2.1 *test (or testing), n*—a procedure performed on an object (or set of nominally identical objects) using specified equipment that produces data unique to the object (or set).

5.2.1.1 *Discussion*—Test data are used to evaluate or model selected properties or characteristics of the object (or set of objects). The scope of testing depends on the decisions to be made for any program, and sampling and replication plans (see definitions below) need to be specified for a complete program description.

5.2.2 *test tire, n*—a tire used in a test.

5.2.3 *test program, n*—an ordered series of tests grouped together using a predefined plan.

5.2.3.1 *Discussion*—A test program may include multiple test repetitions over an extended time period.

5.2.4 *test tire set, n*—one or more test tires as required by the test equipment or procedure, to perform a test, thereby producing a single test result.

5.2.4.1 *Discussion*—The four nominally identical tires required for vehicle stopping distance testing constitute a test tire set. In the discussion below where the test tire is mentioned, it is assumed that test tire set may be substituted for test tire, if a test tire set is required for the testing.

5.2.5 *candidate tire, n*—a test tire that is part of a test program.

5.2.5.1 *Discussion*—The term "candidate object" may be used in the same sense as *candidate tire*.

5.2.6 *candidate tire set*—a set of candidate tires.

5.3 Tire testing may be divided into two major categories:

5.3.1 *local testing, n*—testing conducted at one laboratory or test site for the purpose of comparing a number of candidate tires for selected characteristic properties.

5.3.1.1 *Discussion*—A tire manufacturer's internal development programs and proving ground testing conducted by a contract testing organization to compare commercial market tires are two examples of local testing.

5.3.2 *global testing, n*—testing conducted at two or more laboratories or test sites for the purpose of comparing candidate tire performance at each location for selected characteristic properties.

5.3.2.1 *Discussion*—Producer-user testing or interlaboratory comparisons for such properties as rolling resistance, endurance, or high speed dynamometer wheel performance are examples of global testing.

5.4 *sample, n*—a selected number of *n* test objects that accurately represent the lot or population of interest.

5.4.1 *Discussion*—A lot is a finite number of objects such as a limited period of tire production at a given facility or a selected number of tires of a particular commercial market type. A population is the collection (or potential collection) of all objects produced by a given process or operation.

5.5 *sampling, v*—the act of selecting samples.

5.5.1 *Discussion*—The primary purpose of sampling is the reduction of random production process variation. See Annex A1 for details.

5.6 *replicate, n*—either *(1)* an individual test object from a sample of *n* objects or *(2)* one of *m* individual test values for a test object.

5.6.1 *Discussion*—Each test object of a set of replicates is nominally identical to all other objects from that particular source. Nominally identical implies that in long run testing all objects would give essentially identical average test values.

5.7 *replication, v*—the act of selecting and testing a number of replicates.

5.7.1 *Discussion*—The primary purpose of replication is the reduction of random measurement variation. See Annex A1 (A1.3.6) for additional discussion on types of replication.

6. Reference Tire Concepts and Terms

6.1 In this section a basic term, reference tire, is defined. A number of terms, each describing a special type of reference tire, are derived from the basic term. Reference tires usually have special characteristics unique to a particular test program. However, for some testing programs the same reference tire may be used for more than one purpose.

6.1.1 *reference tire, n*—a special tire included in a test program; the test results for this tire have significance as a base value or internal benchmark.

6.1.2 There are two types of reference tires or objects that may be used in any test program.

6.1.2.1 *Type 1 (reference tire), n*—tires subject to production, composition, and often, performance specifications; they are designed to have minimal variation and to be stable in their characteristic properties for an extended period of time.

6.1.2.2 *Type 2 (reference tire), n*—tires appropriately selected from a lot by a process that ensures minimal variation characteristic properties for the duration of any test program.

6.1.3 *Discussion*—Type 2 reference tires may be selected on an ad hoc basis and when the test program is complete they are no longer considered as reference objects.

6.1.4 *control tire, n*—a reference tire used in a specified manner throughout a test program.

6.1.4.1 *Discussion*—A control tire may be of either type and typical tire use is the reference (control) tire in Practice F 1650 that provides algorithms for correcting (adjusting) test data for bias trend variations (See Practice F 1650 and Annex 1).

6.1.5 *surface monitoring tire, n*—a reference tire used to evaluate changes in a test surface over a selected time period.

6.1.6 *standard reference test tire (SRTT), n*—a tire that meets the requirements of Specification E 1136, commonly used as control tire or a surface monitoring tire.

6.1.6.1 *Discussion*—This is a Type 1 reference tire.

6.1.7 *witness tire, n*—a reference tire with an extended period of stability for specified characteristic properties.

6.1.7.1 *Discussion*—A Type 1 reference tire is typical for this application.

6.1.8 *master set, n*—a selected group of witness tires, each different test response characteristics to provide a range of values for the measured property or properties.

6.1.8.1 *Discussion*—A master set is frequently tested to determine if a test device is functioning in a normal or intended manner. If certain known or expected relationships are not found among the witness tires constituting the set, remedial action is required for the testing equipment. Master sets are frequently used for global testing.

6.1.9 *test matrix, n*—a group of candidate tires usually specified reference tires; all tests are normally conducted in one test program.

6.1.9.1 *Discussion*—A test matrix may be used in either a local or global test program. See also *candidate tire set*.

6.1.10 *calibration tire, n*—a witness tire designed to provide a fixed or known test value for selected properties.

6.1.10.1 *Discussion*—Calibration tire test results can be used as standard values to determine acceptability of laboratory or test site performance. If a specified performance level is not found, certain instrument adjustments may be made to compensate for unavoidable biases in interlaboratory or between-site programs.

7. Evaluating Testing Precision

7.1 As indicated in Annex A1, there are two categories of variation: production process and measurement. Each of these may in turn have two subclassifications: basis deviations and random deviations. The potential effect of all these sources can exert a profound influence on the variability of test data. The presence of these sources is the rationale for using reference tires and for designing comprehensive testing programs with appropriate replication to reduce the effect of such variations.

7.2 *Evaluating Precision*—Special programs to evaluate the magnitude of variability for any routine or special test operations are part of the effort to reduce variability and improve test precision. Precision is defined in Practice F 1082 as "a measurement (testing) concept that expresses the ability to generate test results that agree with each other in absolute magnitude." The parenthetical word "testing" is added to this definition for this purposes of this practice to indicate that is is the overall testing process, which includes sampling and replication, that should be considered when discussing precision.

7.2.1 For local testing, this action usually consists of appropriate sampling and replication plans and the evaluation of "test-to-test" variation for candidate tires. With a "test-to-test" standard deviation (or variance) obtained under the appropriate conditions, decisions on statistical (and technical) differences between candidate tires can be made for a program at any specific location. For global testing, programs can be organized to evaluate another "test-to-test" standard deviation, where this now applies to between-lab as well as to between-test comparisons.

7.3 *Repeatability and Reproducibility*:

7.3.1 The terms repeatability and reproducibility are frequently used when discussing testing and the results of testing programs. Some interpretations of these terms are different than the standard definitions given in Practice F 1082.

7.3.1.1 *repeatability, n*—an established value, below which the absolute difference between two "within-laboratory" or "within test site" test results may be expected to lie, with a specified probability (Practice F 1082).

7.3.2 *Discussion*—The two test results are obtained with the same method on nominally identical test materials under the same conditions (same operator, apparatus, laboratory, location, and specified time period), and in the absence of other indications, the specified probability is 0.95 (that is, 95 %). The established value also may be called a "critical difference."

7.3.2.1 *reproducibility, n*—an established value, below which the absolute difference between two "between-laboratory" or "between test site" test results may be expected to lie, with a specified probability (Practice F 1082).

7.3.3 *Discussion*—The two test results are obtained with the same method on nominally identical test materials under different conditions (different laboratories, locations, operators, apparatus, and in a specified time period), and in the absence of other indications, the specified probability is 0.95 (that is, 95 %). The essential characteristic of reproducibility is the variability of test results among typical laboratories or test sites.

7.3.4 Both repeatability and reproducibility are to some degree generic in their definition. Additional information must be supplied before the terms can be used without ambiguity. The most important issue is the between–test result time period or frequency; it must be specified. What constitutes a test result must be defined. Both of these are addressed in Practice F 1082. Other details on testing are also needed. It is important to emphasize two details about repeatability and reproducibility; (*1*) both are statistical parameters; defined as $2.83\times S$, where S is the standard deviation for either parameter measured in a specified way as outlined in Practice F 1082 and (*2*) both parameters indicate precision in an inverse manner; high precision equals small values for either parameter.

NOTE 1—One source of confusion is the use of the words repeatability and reproducibility alone to indicate a desired or high level of precision. "A test has repeatability" is an inappropriate use of the word repeatability if consistency with Practice F 1082 is desired. "A test has good (or low) repeatability" is appropriate usage.

7.4 *Process and Measurement Variation*:

7.4.1 Annex A1 contains a brief section (A1.4) on evaluating process and test measurement variance. This can be done rather easily for a non destructive test such as tire wet traction testing or rolling resistance testing. A more detailed analysis on this topic is beyond the scope of this practice in its present format.

8. Keywords

8.1 control tire; monitoring tire; reference tire; repeatability; reproducibility; test matrix; witness tire

ANNEX

(Mandatory Information)

A1. STATISTICAL MODEL FOR TEST MEASUREMENT

A1.1 Background

A1.1.1 The purpose of this annex is to present some of the concepts and definitions used and implied when test measurement variation is discussed. Using a mathematical model format improves understanding and more clearly demonstrates how the variation concepts relate to each other. In the annex, some of the words or terms are given a specific definition; some are informally defined by the context of their use.

A1.1.2 All measurement values are perturbed to some degree by a "system-of-causes" that produces error or variation in the measured parameter. There are two general variation categories for any system:

A1.1.2.1 *Measurement Variation*—deviations in the operation of devices or machines that evaluate certain properties for any class of objects or material; these deviations perturb the observed values for these properties.

A1.1.2.2 *Production variation*—deviations in certain properties that are (*1*) inherent in the production process that produces the different classes of objects or materials being tested or (*2*) acquired deviations (storage or conditioning effects) after such processes are complete.

A1.1.2.3 *Discussion*—Certain types of variation may be inherent in any particular realization of a test device, that is, independent of the test device operation.

A1.1.3 The system-of-causes is defined by the scope and organization of any testing program and by the replication and sampling operations that are part of the program. These systems can vary from simple to very complex. The production process can be (*1*) the ordinary operation of a manufacturing facility or (*2*) some smaller processing or other operation that produces a material or class of objects for testing.

A1.2 General Model

A1.2.1 For any established "system-of-causes," each measurement, y (i), can be represented as a linear additive combination of fixed or variable (mathematical) terms as indicated by Eq A1.1. Each of these terms is an individual component of variation and the sum of all components is equal to the total variation observed in the measurement operation. The equation applies to any brief time period of testing for a standardized test procedure. All participants test a number of classes of objects (each class having a number of individual nominally identical objects) or different materials, drawn from a lot of some specified uniformity, employ the same type of apparatus, use skilled operators, and are conducting testing in a typical environment.

$$y(i) = \mu(o) + \mu(j) + \Sigma(b) + \Sigma(e) + \Sigma(\beta) + \Sigma(\epsilon) \quad (A1.1)$$

where:

$y(i)$ = a measurement value, at time (i), using specified equipment and operators, at laboratory (q),

$\mu(o)$ = a general or constant term or mean value, unique to the type of test being used,

$\mu(j)$ = a constant term (mean value), unique to material or object class (j),

$\Sigma(b)$ = the (algebraic) sum of some number of individual bias deviations in the process that produced material or object class (j),

$\Sigma(e)$ = the (algebraic) sum of some number of individual random deviations in the process that produced material or object class (j),

$\Sigma(\beta)$ = the (algebraic) sum of some number of individual bias deviations, for measurement (i), generated by the measurement system, and

$\Sigma(\epsilon)$ = the (algebraic) sum of the number if individual random deviations, for measurement (i), generated by the measurement system.

A1.2.2 Eq A1.1 identifies three main sources of generic variation components: (*1*) constant terms (population mean values); (*2*) bias deviation terms, and (*3*) random deviation terms. These three are discussed in detail in succeeding sections.

A1.3 Specific Model Format

A1.3.1 A more useful format is obtained when Eq A1.1 is expressed using Eq A1.2, where the summations are replaced by a series of typical individual terms appropriate to interlaboratory testing on a number of different objects or materials, for a particular time period sufficient to complete the testing. This permits greater insight into the model and how it relates to real testing situations.

$$y(i) = \mu(o) + \mu(j) + \Sigma b + \Sigma e + \beta(L) + \beta E) + \beta(OP) + \epsilon(E) + \epsilon(OP) \quad (A1.2)$$

where:

$\beta(L)$ = a bias deviation term unique to laboratory (q),

$\beta(E)$ = a bias deviation term unique to the specific equipment or machine,

$\beta(OP)$ = a bias deviation term unique to the operator (s) conducting the test,

$\epsilon(E)$ = a random deviation inherent in the use of the specific equipment, and

$\epsilon(OP)$ = a random deviation inherent in operator's technique.

Other types of testing perturbations not included in Eq A1.2 may exist, such as bias and random components due to temperature and other factors such as the time of the year that testing is conducted.

A1.3.2 *The $\mu(o) + \mu(j)$ Terms*—In the absence of bias or random deviations of any kind, a set of objects would have individual measured test values given by the sum of the two terms, $\mu(o) + \mu(j)$. The term $\mu(o)$ would be unique to the test employed and each candidate would be characterized by the value of $\mu(j)$, which would produce a varying value for the sum

[$\mu(o) + \mu(j)$] for comparisons among any number of candidates. The sum [$\mu(o) + \mu(j)$] would be the "true" test value for any candidate, that is, without error or variation of any sort.

A1.3.3 *The Production Terms* $\Sigma(b) + \Sigma(e)$—There will always be some bias and random variation in the candidate test objects produced by the process that generates them. These usually unknown number of bias and random variations are designated by $\Sigma(b) + \Sigma(e)$. Special production operations and other precautions can frequently be employed to reduce this variation to a level where it is substantially less than the test variation. In some testing programs this production variation is assumed to be zero. However, if such special precautions are not taken or for highly accurate evaluation programs, the production process variation must be taken into account. This is accomplished by appropriate program organization and the use of plans that allow for replicate sampling of the lots of objects for each candidate set.

A1.3.4 *The Measurement Bias (β) Terms*—The classic statistical definition of a bias is "the difference between the average measured test results and the accepted reference value (true value); it measures in an inverse manner the accuracy of a test", see Practice F 1082. Bias deviations are non-random components and for a series of extended measurements (a long run) the value of bias terms may be either fixed or variable as well as + or - , depending on the system-of-causes. The variable bias terms are typically a non-random finite distribution which in the long run give a non-zero average. Biases or bias deviations are primarily responsible for the difference among laboratories, machines, or test sites.

A1.3.4.1 Bias terms that are fixed under one "system-of-causes" may be variable under another "system-of-causes" and vice versa. As an example, consider the bias terms $\beta(L)$ and $\beta(E)$ which apply to most types of testing. For a particular laboratory (with one test machine) both of these bias terms would be constant or fixed. For a number of test machines, all of the same design in a given laboratory, $\beta(L)$ would be fixed but $\beta(E)$ would be variable, each machine potentially having a unique value. For a measurement system consisting of a number of typical laboratories, each with one machine, both $\beta(L)$ and $\beta(E)$ would be variable for the multilaboratory "system-of-causes," but both $\beta(L)$ and $\beta(E)$ would be fixed or constant for the "system-of-causes" in each laboratory.

A1.3.5 *The Measurement Random (ϵ) Terms*—These are the components that are frequently called error. Random deviations are + or - values that have an expected mean (average) of zero over the long run. The distribution is unimodal. The value of each random term influences the measured y (*i*) value on an individual measurement basis. However in the long run when y (*i*) values are averaged over a substantial number of measurements, the influence of the random terms may be greatly diminished or eliminated depending on the sampling and replication plan, since each term averages out to zero (or approximately zero) and the average y(*i*) is essentially unperturbed. In ordinary testing the magnitude of the individual bias and random components or deviations are not known. Their collective effect influences each measured y (*i*) value and this collective effect is what is normally evaluated in variance testing.

A1.3.6 *Test Replication*—For any given test, there are three general types of sample replication which apply to the number of objects tested; (*1*) Type 1 sample replication (*m*), using the same test object with 1 to *m* repeated tests, (*2*) Type 2 sample replication (*n,1*), using *n* test objects, each object being tested one time, and (*3*) Type 3 sample replication (*n,m*), using *n* test objects, each object being tested *m* times.

For Type 1, the sample size is 1, with *m* replicates; for Type 2 and 3 the sample size is *n*, also with *m* replicates. The scope of the sampling and replication plan needs to be clearly defined for any testing program. Replication Types 1 (with *m* tests) and 3 may be used for non destructive testing, while Type 2 is the only type available for multi-sample destructive testing. Type 3 testing reduces the influence of the production random variation as well as the random measurement variation.

A1.3.6.1 Replicated testing of any type with only a few replicates (where *n* and *m* jointly or each equal less than ten gives a test result average value, $Y(n, m < 10)$, as indicated by Eq A1.3, where the appearance of $\Sigma(e)$ and $\Sigma(\epsilon)$ indicates that these sums are not equal to zero. Usually $\Sigma(\epsilon)$ and $\Sigma(e)$ are much less than $\Sigma(b)$ and $\Sigma(\beta)$.

$$Y(n,m < 10) = \mu(o) + \mu(j) + \Sigma(b) + \Sigma(e) + \Sigma(\beta) + \Sigma(\epsilon) \quad (A1.3)$$

A1.3.6.2 Highly replicated testing (ten or more measurements for both *n* and *m*) reduces the perturbation of the random deviations to near zero. Thus the test result average $Y(n,m>10)$, is given by Eq A1.4, which is perturbed by only bias components.

$$Y(n,m>10) = \mu(o) + \mu(j) + \Sigma(b) + \Sigma(\beta) \quad (A1.4)$$

A1.3.6.3 Eq A1.4 shows that ordinary highly replicated testing (usually Type 3) does not approximate the "true value" for any candidate if any production or measurement system bias deviations exist. The tester ordinarily does not not know of the potential sources of this inherent process and measurement bias variation. Therefore for future discussion in this annex, no individual assignment of variation components is made. These terms remain in their generalized format.

A1.3.7 *New Term, M(j)*—With highly replicated programs (both production and testing replication) the average values obtained for any candidate in any program are estimates or very close approximations to the value of a new combined term as given by Eq A1.5.

$$M(j) = [\, \mu(o) + \Sigma(b) + \Sigma(\beta) \,] + \mu(j) \quad (A1.5)$$

M (*j*) is the mean value for the objects of candidate set (*j*) for the test being used, for laboratory or test site (*q*), for the specific equipment and operators used during the existing time period. It contains bias components or potential bias components for all of these conditions. If all biases are fixed for any given program, the three terms in the bracket can be considered as a constant and the average test value for candidates varies across the number of candidates because of the varying value of $\mu(j)$. If there are variable biases, then both $\mu(j)$ and the biases influence the average value for any candidate.

A1.4 Evaluating Process and Measurement Variance

A1.4.1 Eq A1.1 may be used to illustrate how the variance of individual measurements, *y* (*i*), may be related to the terms or components of the equation. Recall that μ(*o*) and μ(*j*) are constants, Σ (*b*) and Σ (*e*) refer to the sum of bias and random components, respectively, for the production process and Σ (β) and Σ (ϵ) refer to the sum of bias and random components, respectively, for the test measurement operation. Since the tester ordinarily does not know the magnitude of the individual components, the equation can be simplified by combining the bias and random components for both sources.

$$y(i) = \mu(o) + \mu(j) + \Sigma(b,e) + \Sigma(\beta, \epsilon) \quad \text{(A1.6)}$$

where:

$\Sigma(b, e)$ = sum of bias and random components for the production process, and

$\Sigma(\beta, \epsilon)$ = sum of bias and random components for the measurement operation.

A1.4.2 The variance of any individual measurement *y* (*i*), designated by Var [*y* (*i*)], is given in Eq A1.7

$$\text{Var}[y(i)] = [\Sigma \text{Var}(b,e)] + [\Sigma \text{Var}(\beta,\epsilon)] \quad \text{(A1.7)}$$

where:

$[\Sigma \text{Var}(b,e)]$ = a variance that is the sum of variances of individual bias and random variances for the production process, and

$[\Sigma \text{Var}(\beta,\epsilon)]$ = a variance that is the sum of variances of individual bias random variances for the measurement operation.

Eq A1.7 can be written in simplified format as indicated in Eq A1.8, using the conventional symbol S^2 for the variance.

$$S^2(\text{tot}) = S^2(p) + S^2(m) \quad \text{(A1.8)}$$

where:

$S^2(tot)$ = total variance among measured objects in a test program, and

$S^2(p)$ = variance due to production process, and

$S^2(m)$ = variance due to measurement operation.

All three variance components can be evaluated for a nondestructive test where any sample may be tested more than one time. Table A1.1 will help in illustrating this for a typical testing scenario. There are (*k*) candidates tested, each candidate has a sample of four (*n* = 4) and each of the four sample objects is tested two times (*m* = 2). Each pair of *y* (*i j*) -values constitutes a cell in the table.

TABLE A1.1

Candidate Object	Sample No. 1	2	3	4
A	y11, y12	y21, y22	y31, y32	y41, y42
B	etc.	...	...	...
...	...	...	...	...
k	...	...	...	...

A1.4.3 There are (*k*) *8 individual test values and the variance for these values is S^2(tot). The variance $S^2(m)$ is evaluated by taking the variance for each cell in the table (each cell has 1 DF) and pooling this across cells for all candidates. The variance $S^2(p)$ is evaluated by difference as given in Eq A1.9.

$$S^2(p) = S^2(\text{tot}) - S^2(m) \quad \text{(A1.9)}$$

This approach to production process and test measurement variance evaluation assumes that the replicate testing variance (within cell) is equal for all candidates and thus the value of $S^2(p)$ as obtained from this analysis is a collective value representing all candidates.