APPENDIX C

TERMINOLOGY OF THE GUIDE TO THE EXPRESSION OF UNCERTAINTY IN MEASUREMENT (GUM)

This thesis deals with *accuracy*. It is then imperative that a rigorous terminology be followed throughout. Reporting the result of a measurement of a physical quantity requires that some quantitative indication of the quality of the result be given so that reliability of the latter can be assessed. Besides, it is necessary that no confusion arises when terms such as *uncertainty*, *error*, *precision* or *accuracy* are used for instance. In the years following 1978, a great deal of effort has been made by the Bureau International des Poids et Mesures (BIPM) in order to obtain an international consensus on the expression of the uncertainty in measurement along with establishment of a proper vocabulary of metrological and statistical terminology. The task of developing a detailed guide fell to the International Organisation for Standardization (ISO), and in 1993 they came out with the *Guide to the expression of uncertainty in measurement*, or GUM [ISO93]. Throughout the thesis we have followed narrowly the recommendations and the terminology as given in GUM. This appendix aims at highlighting some of the more important definitions found in the guide; these are presented as (nearly) literally given in GUM, with some extra details added for better clarity.

C.1 Measurand

The measurand is the particular quantity subject to measurement. For example: the vapour pressure of a given sample of water at 20ºC. Note that the specification of a measurand may require statements about quantities such as time, temperature, pressure, etc.

C.2 Conventional true value

The conventional true value of a quantity (measurand) is the value attributed to that particular quantity and accepted, sometimes by convention, as having an uncertainty (see further below for the definition of uncertainty) appropriate for a given purpose. For example, at a given location, the value assigned to the quantity realized by a reference standard may be taken as a conventional true value. Another example would be the 1986 CODATA (Committee on Data for Science and Technology) recommended value for the Avogadro constant: 6.0221367×10^{23} mol^{-1}.

"Conventional true value" is sometimes called "assigned value", "best estimate" of the value, "conventional value" or "reference value". The term "certified value" may be added to the list. Frequently, a number of results of measurements of a quantity is used to establish a conventional true value.

C.3 Accuracy and precision

Accuracy and precision are two terms often misunderstood and confused. The accuracy of a measurement is the closeness of the agreement between the results of a measurement and a conventional true value of the measurand; in other words, accuracy is how close to the accepted value a measurement lies. In contrast, precision is a measurement of how closely the analytical results can be duplicated; thus precision measures how far from the mean of replicate measurements a particular measurement lies, and it can be reported as a standard deviation. Accuracy is a qualitative concept; the term precision should not be used for accuracy.

If a conventional true value is represented as a bull's eye on a target, a group of guesses or measurements represented by closely grouped points have a high degree of precision. If this group is near the centre, it is highly accurate as well. On the other hand, if the points are widely scattered around the centre, the measurements can be said accurate but not highly precise. These situations are depicted in Figure C-1.

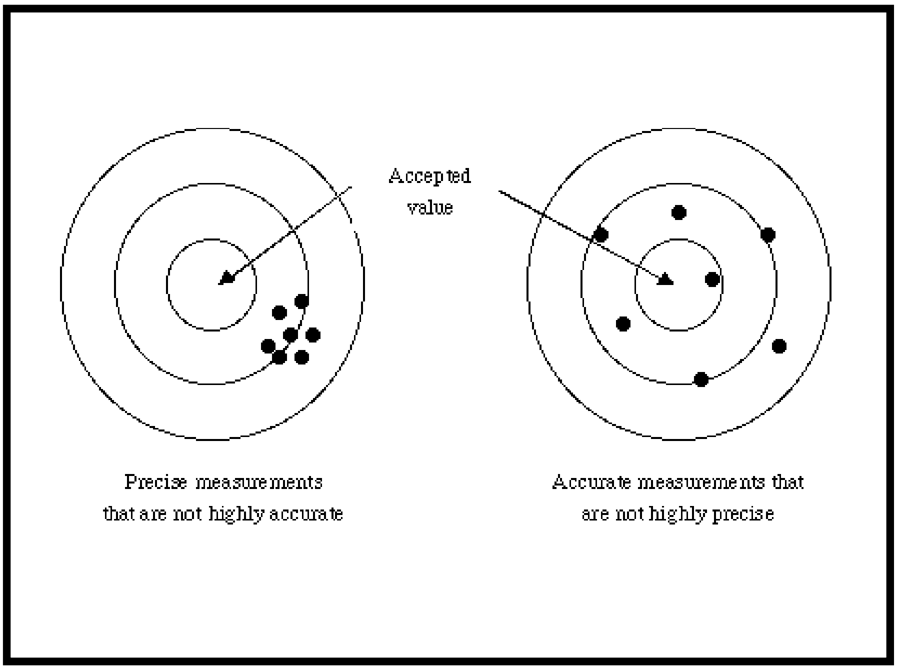

Figure C-1 *Distinction between accuracy and precision.*

C.4 Basic statistical terms and concepts

C.4.1. Expectation — mean value of random variable

The expectation of the random variable z, denoted by μ_z, and which is also termed the expected value or the mean of z, is given by:

$$\mu_z = \int z p(z) dz \,, \qquad \text{(C-1)}$$

where $p(z)$ is the probability density function of the random variable z. From the definition of $p(z)$

$$\int p(z) dz = 1 \,. \qquad \text{(C-2)}$$

This expectation is estimated statistically by $\bar{z}$, the arithmetic mean or average of *n* independent observations z_i of the random variable *z* (measurand), the probability density function of which is *p(z)*:

$$\bar{z} = \frac{1}{n}\sum_{i=1}^{n} z_i \ . \qquad \text{(C-3)}$$

C.4.2. Variance of random variable

The variance of a random variable is the expectation of its quadratic deviation about its expectation. Thus the variance of random variable *z* with probability density function *p(z)* is given by:

$$\sigma^2(z) = \int (z - \mu_z)^2 p(z) dz \ . \qquad \text{(C-4)}$$

The variance may be estimated by:

$$u^2(z_i) = \frac{1}{n-1}\sum_{i=1}^{n} (z_i - \bar{z})^2 \ , \qquad \text{(C-5)}$$

where the z_i are *n* independent observations of *z*. Note that the factor *n-1* arises from the correlation between z_i and $\bar{z}$ and reflects the fact that there are only *n-1* independent items in the set $\{z_i - \bar{z}\}$.

C.4.3. Variance of arithmetic mean

The variance of arithmetic mean or average of the observations, rather than the variance of the individual observations (variance of random variable), is sometimes used to measure the uncertainty of a measurement result. The variance of a variable *z* should be carefully distinguished from the variance of the mean $\bar{z}$. The variance of the arithmetic mean of a series of *n* independent observations z_i of *z* is estimated by the experimental variance of the mean:

$$u_m^2 = \frac{u^2(z_i)}{n} = \frac{1}{n(n-1)}\sum_{i=1}^{n} (z_i - \bar{z})^2 \ . \qquad \text{(C-6)}$$

C.4.4. Standard deviation of random variable

The standard deviation for a series of n measurements of the same random variable z (measurand) is the positive square root of the variance, i.e. u, and it characterizes the dispersion of the results.

C.4.5. Standard deviation of arithmetic mean

The value u_m (square root of the variance of the mean as expressed further above) is an estimate of the standard deviation of the distribution of the arithmetic mean $\bar{z}$, and it is called the standard deviation of the mean.

C.5 Uncertainty

The word "uncertainty" means doubt, and thus in its broadest sense "uncertainty of measurement" means doubt about the validity of the result of a measurement. Because of the lack of different words for this *general concept* of uncertainty and the specific quantities that provide *quantitative measures* of the concept, e.g. the standard deviation, it is necessary to use the word "uncertainty" in these two different senses.

C.5.1. Uncertainty of measurement

Uncertainty of measurement is the parameter, associated with the result of a measurement, that characterizes the dispersion of the values that could reasonably be attributed to the measurand. The parameter may be, for example, a standard deviation (or a given multiple of it), or the half-width of an interval having a stated level of confidence.

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 also can be characterized by standard deviations, are evaluated from assumed probability distributions based on experience or other information.

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.

C.5.2. Standard uncertainty

The standard uncertainty is when the uncertainty of the result of a measurement is expressed as a standard deviation.

C.5.3. Categories of uncertainty

The uncertainty in the result of a measurement generally consists of several components which may be grouped into two categories, denoted Type A and Type B, according to the way in which their numerical value is estimated. The purpose of the Type A and Type B classification is to indicate two different ways of evaluating uncertainty components and is for convenience of discussion only; the classification is not meant to indicate that there is any difference in the nature of the component resulting from the two types of evaluation. Both types are based on *probability distributions*, and the uncertainty components resulting from either type are quantified by variances or standard deviations.

C.5.3.1. Uncertainty: Type A

Components that are evaluated by statistical methods (statistical analysis of series of observations). They are characterized by the estimated variances or standard deviations. The estimated variance u^2 characterizing an uncertainty component obtained from a Type A evaluation is calculated from series of repeated observations and is the familiar statistically estimated variance as given by equation (C-5).

C.5.3.2. Uncertainty: Type B

Components whose method of evaluation of uncertainty is by means other that the statistical analysis of series of observations. For an uncertainty component obtained from a Type B evaluation, the estimated variances (or standard deviations) are evaluated using available knowledge, and may be also denoted u^2, which may be considered as approximations to the corresponding variances, the existence of which is assumed. These quantities u^2 may be treated like variances.

C.5.4. Combined standard uncertainty

The standard uncertainty of the result of a measurement, when that result is obtained from the values of a number of other quantities, is termed combined standard uncertainty and denoted u_c. It is the estimated standard deviation associated with the result and is equal to the positive square root of the combined variance obtained from all variance components summed (law of propagation of uncertainty — sum in quadrature).

C.5.5. Expanded uncertainty

The expanded uncertainty can be said the 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. The fraction may be viewed as the coverage probability or level of confidence of the interval.

When it is necessary to multiply the combined uncertainty by a factor, called coverage factor, to obtain an overall uncertainty, the multiplying (coverage) factor used must always be stated. A coverage factor, usually denoted k, is typically in the range 2 to 3.

C.6 Error of measurement

The error of measurement corresponds to the result of a measurement minus a (conventional or accepted) true value of the measurand.

C.6.1. Relative error

The relative error is the error of measurement divided by a (conventional or accepted) true value of the measurand (usually expressed in percentage).

C.6.2. Random error

The random error is the result of a measurement minus the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions. Because only a finite number of measurements can be made, it is possible to determine only an estimate of random error. Note that the random error is equal to error minus systematic error.

C.6.3. Systematic error

The systematic error is equal to the mean that would result from an infinite number of measurements of the same measurand carried out under repeatability conditions minus a (conventional or accepted) true value of the measurand. Like true value, the systematic error and its causes cannot be completely known. Note that the systematic error is equal to error minus random error.

C.6.4. Correction factor

The numerical factor by which the uncorrected result of a measurement is multiplied to compensate for systematic error is called correction factor. Since the systematic error cannot be known perfectly, the compensation cannot be complete.

C.7 Repeatability and reproducibility of results of measurements

C.7.1. Repeatability

The repeatability of results of measurements is the closeness of the agreement between the results of successive measurements of the same measurand carried out under the same conditions of measurement. The repeatability conditions include: the same measurement procedure; the same observer; the same measuring instrument, used under the same conditions; the same location; repetition over a short period of time. Repeatability may be expressed quantitatively in terms of the dispersion characteristics of the results.

C.7.2. Reproducibility

The reproducibility of results of measurements is the closeness of the agreement between the results of measurements of the same measurand carried out under changed conditions of measurement. A valid statement of reproducibility requires specification of the conditions changed. The latter may include: principle of measurement; method of measurement; observer; measuring instrument; reference standard; location; conditions of use; time. Reproducibility may be expressed quantitatively in terms of the dispersion characteristics of the results.