Int. J. Metrol. Qual. Eng. **3**, 155–159 (2012)
© EDP Sciences 2013
DOI: 10.1051/ijmqe/2012024

Corrections and input quantities in measurement models

F. Pavese*

Torino, Italy

Received: 15 October 2012 / Accepted: 28 October 2012

Abstract. The *Guide to the expression of uncertainty in measurement* (GUM) requires, "that the results of a measurement have been corrected for all recognised significant systematic effects and that every effort has been made to identify such effects", before the issue of evaluating their uncertainty is tackled. In addition, the GUM also requires that a model is used, expressing the functional dependence of the measurand on identified input quantities. The paper discusses the inter-relations between the so-defined two sets of influence quantities and proposes to unify their treatment.

Keywords: GUM; correction; influence quantity; model; systematic effect

1 Introduction

The *Guide to the expression of uncertainty in measurement* (GUM) [1] introduced since 1995 methods to be used in the field of metrology for evaluating the measurement uncertainty. According to this Guide, "*it is assumed that the results of a measurement have been corrected for all recognised significant systematic effects and that every effort has been made to identify such effects*" (3.2.4), before the issue of evaluating their uncertainty is tackled. In addition, in the GUM it is also required that a model is used, expressing the functional dependence of the measurand on the identified input quantities X_i: (GUM, 4.1.1) "*In most cases, a measurand Y is not measured directly, but is determined from N other quantities $X_1, X_2, \ldots, X_N$ through a functional relationship f*:

$$Y = f(X_1, X_2, \ldots, X_N)''$$

(GUM, 4.1.2) "*The input quantities $X_1, X_2, \ldots, X_N$ upon which the output quantity Y depends may themselves be viewed as measurands and may themselves depend on other quantities, including corrections and correction factors for systematic effects, thereby leading to a complicated functional relationship f that may never be written down explicitly. Further, f may be determined experimentally ... or exists only as an algorithm that must be evaluated numerically. The function f as it appears in this Guide is to be interpreted in this broader context, in particular as that function which contains every quantity,* ***including all corrections and correction factors*** *that can contribute a significant component of uncertainty to the measurement result*" (boldface added here).

The function may remain implicit $f(Y, X_i) = 0$, where the effect of each X_i on Y is obtained through a Type A variational method for the determination of the "*sensitivity coefficients*", or based on Type B information; or, the function can be an explicit mathematical expression, whose values of the parameters are determined experimentally, or via Type B information.

The wished effects of correction for systematic effects, as expressed in the GUM, are that: "*it is assumed that, after correction, the expectation or expected value of the error arising from a systematic effect is zero*" (3.2.3); this outcome "*... is based on the concept that there is no inherent difference between an uncertainty component arising from a random effect and one from a correction for a systematic effect*" (E.1.1), because "*... this Guide treats uncertainty components arising from random effects and from corrections for systematic effects in exactly the same way in the evaluation of the uncertainty of the result of a measurement.*" (E.3).

However, correction for systematic effects may actually be affected by limitations, as also recognised in limited cases in the GUM: "*It is now widely recognized that, when all of the known or suspected components of error have been evaluated and the appropriate corrections have been applied, there still remains an uncertainty about the correctness of the stated result, that is, a doubt about how well the result of the measurement represents the value of the quantity being measured*" (0.2); "*The result of a measurement after correction for recognized systematic effects is still only an estimate of the [true] value of the measurand because of the uncertainty arising from random effects and from imperfect correction of the result for systematic effects*" (3.3.1); "*A corrected measurement result is not the value of the measurand – that is, it is in error – because of imperfect measurement of the realized quantity due to random variations of the observations (random effects), inadequate determination of the corrections for systematic effects, and incomplete knowledge of certain physical phenomena (also systematic effects). Neither the value of the realized quantity nor the value of the measurand can ever be known exactly; all that can be known is their estimated values.*" (D.4).

* Correspondence: frpavese@gmail.com

Article published by EDP Sciences

https://doi.org/10.1051/ijmqe/2012024 Published online by Cambridge University Press

As a consequence, a live debate is going on about the way to comply with the GUM requirements, and, more in general, to comply with the need of performing corrections for known biases, or for biases estimated at various levels of confidence – as occurring also in the testing field.

2 Problems to tackle about corrections

In fact, the above caveats should concern not only the uncertainty, but also the value of the correction, since a model is likely to be imperfect for several reasons:

- non-uniqueness of the measurand (epistemic uncertainty, intrinsic);
- incomplete list of influence quantities (epistemic uncertainty);
- imperfect knowledge of the input quantities (re model functional form and values of its parameters).

In addition, in tackling the problem, one is often confronted with issues arising from practical difficulties, especially when:

(a) applying a significant recognised correction with large uncertainty;
(b) handling uncertainty for uncorrected results.

Concerning issue (a), two conditions have to be noted: (i) to be "recognised" (be known), (ii) to be "significant", also embedded in the GUM statement: "*In some cases, . . . although the uncertainty has been evaluated, it may be ignored if its contribution to the combined standard uncertainty of the measurement result is insignificant. If the value of the correction itself is insignificant relative to the combined standard uncertainty, it too may be ignored*" (3.4.4). However, e.g., CITAC [2] recommends instead correcting also for non-significant effects.

As to issue (b), this use is widespread. Actually, in principle, the corrections are fundamental good practice in science, because they are to be considered as an integral part of the process of sufficiently understanding and modelling an experiment. They are as essential as the practice of replicating the measurements. In practice, there are several conditions under which the correction of the measured value is not desirable, and this is not explicitly deprecated in the GUM but allowed in special cases (see Sect. 3).

GUM's way out from the difficulties in applying a correction is the following:

- "*The error arising from imperfect compensation of a systematic effect cannot be exactly known*" (3.2.3. Note), . . .

but specifying that

- "*The uncertainty of a correction for a known systematic effect may in some cases be obtained by a Type A evaluation while in other cases by a Type B evaluation, as may the uncertainty characterizing a random effect.*"

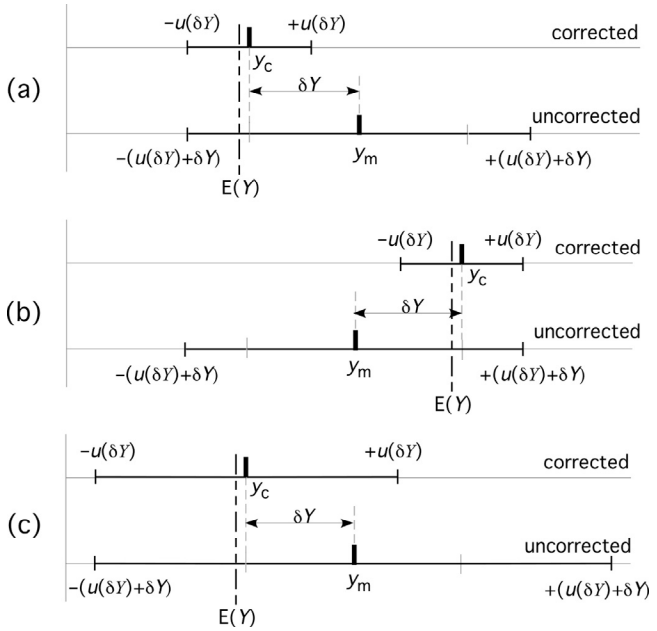


Fig. 1. Correction/no-correction process and outcomes: $E(Y)$ = aimed Y value; y_c = value after application of the estimated correction $\delta Y \approx \Delta$ to the measured value y_m; $u(\delta Y)$ = uncertainty of estimated correction δY. It can be $y_c - y_m > u(\delta Y)$.

In fact, fixed effects are assumed in the GUM to be fully taken into account by the corrections (zero-expectation after correction), so that only random effects apply to the corrected value. One may wonder if this very strong assumption may generally apply.

Let us consider the overall situation, depicted in Figure 1, where Y is measured directly and Δ is the correction, whose value is $\Delta = E(Y) - y_m$, with y_m being the measured value and $E(Y)$ the unknown expectation[1]. In Figure 1 ("corrected"), (a) is for a negative correction with a small uncertainty; (b) is for a positive correction with a small uncertainty; (c) is as (a) for a large uncertainty. According to the GUM assumption, $y_c \equiv E(Y)$ (i.e., $\delta Y \equiv \Delta$).

A peculiar interesting situation occurs when $\delta Y = 0$ (sometimes equivalent to $\Delta = 0$) that can have two commonly used meanings: (1) a correction applied to the measured data of value zero and of a nonzero associated uncertainty; (2) measured data are used directly, without correction, and a method is adopted for evaluation of the resulting additional uncertainty arising from that decision. See [3] for a full discussion on the differences arising from the two different ways of intending the fact that in both cases the measured values are effectively used, namely with respect to the reliability of the resulting estimate of the value assigned to the mean value with respect to an actual 'true value'.

[1] The uncertainty of y_m is taken to be zero in Figures 1 and 2, since the value would not alter the comparison of the "corrected" and "uncorrected" cases.

https://doi.org/10.1051/ijmqe/2012024 Published online by Cambridge University Press

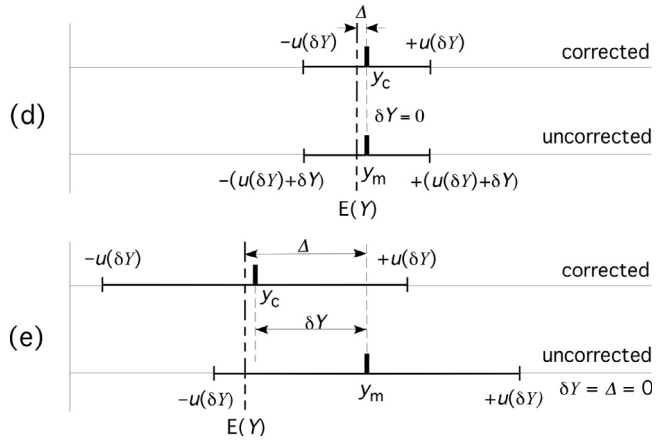


Fig. 2. Correction/no-correction process and outcomes: case when $\delta Y = 0$. Symbols as in Figure 1.

In case (1), setting to zero the value of a correction can arise from basically two distinct reasons:

(1a) the computed correction value is effectively zero within the significant digits needing be used;
(1b) a Type B evaluation brings defining a null correction, $\delta Y := 0$ (in some cases $\Delta := 0$).

The very rarely occurring circumstance (1a) requires, as usual when performing a correction and irrespective to its value, a sufficient confidence in the validity of the computation. Then, it can then be treated in all respect as the case of a value different from zero, either positive or negative: $\delta Y = 0 \pm u(\delta Y)$, a *symmetrical* interval. This case is shown in Figure 2d, on lines labelled "corrected".

Since in the GUM the need is restricted to only significant corrections, one should note that, a zero value (of Δ in this case) also may result from considering a non-zero value as non-significant, according to a decision concerning the significance level.

In case (2), a zero value of the correction occurs quite commonly when the information available is considered insufficient for computing a reliable value of the correction, or even for deciding about its sign. This is more common, e.g., in chemistry: "*Even where a bias is detectable on a related material (such as a reference material), the cause of the bias may not be known, and without a known mechanism, it may be unsafe to apply a correction based on studies of another material*" [4].

In these cases, should one decide to evaluate (i.e., to infer) a correction value, one has to resort (mainly) to Type B evaluation, i.e. not based on statistical methods. This evaluation, may be based, for example, on the fact that the sign of the correction is known [5], bringing to a non-zero correction and to a *non-symmetrical* uncertainty interval. In the limit case, (nearly)[2] no information at all can be considered suitable for the purpose.

[2] Obviously, a total lack of information would prevent from inferring any correction value, and the use of a value zero would also be arbitrary, since no associated uncertainty could be estimated for it. In all instances, here one deals with epistemic ignorance, a case not explicitly treated in the GUM.

However, the assumption that the systematic deviations with positive and negative effects perfectly balance out is a very strong one and rarely justified. In general it increases $u(\delta Y)$ abnormally (see examples in [5]). If the uniform probability distribution (equally probable values) is chosen – meaning full epistemic ignorance – one may wonder how it can be justified to take its centre as the expected value.

3 Why and when corrections may not be applied

Figure 1 ("uncorrected") shows the corresponding different situations when the experimental data are used uncorrected. The full overlap of the uncertainty intervals can be ensured only by taking in consideration the omitted estimate of bias. It should be noted that taking into account an untrusted correction value as an enlargement of uncertainty is much less critical than altering the measured value.

In the GUM there is a warning: "*Occasionally, one may find that a known correction for a systematic effect has not been applied to the reported result of a measurement, but instead an attempt is made to take the effect into account by enlarging the 'uncertainty' assigned to the result. This should be avoided; only in very special circumstances should corrections for known significant systematic effects not be applied to the result of a measurement*" (6.3.1).

Actually, not infrequently the epistemic uncertainty is dominating, making an uncertain correction a wrong decision altering the value of the measurand.

Notably, the mutual recognition arrangement (MRA) in metrology [2] does not advise to change the measured value in case of "*significant unresolved differences*" (SUD) found between laboratories in key comparison (KC) results, but only require to enlarge the uncertainty sufficiently to obtain metrological compatibility [6]. For a discussion of methods intended for only enlarging uncertainty see [3].

When the correction is set to, or is evaluated to be, zero, both methods, of correcting and not correcting the measured data, obviously bring to the same uncertainty u, as shown in Figure 2d.

Basically, when one decides to not correct the measured data, the decision is not generally taken as the 'simplest' way-out. On the contrary, one estimates that a reasonably high confidence on the correction value is not available for altering the measured value. Then, the metrological problem moves to the best estimate of the uncertainty enlargement suitable for taking into account the information that one is aware of being embedded in the systematic effect, and that must be taken into account, but without altering the measured value.

Apparently, this is the same condition of the case when the correction value is computed – or estimated – to be equal to zero. In addition, in the case where one is not trusting the value of the correction to such an extent to

https://doi.org/10.1051/ijmqe/2012024 Published online by Cambridge University Press

prefer using uncorrected data, one may be tempted, especially when the interval of the measured values includes the zero, to assume $\Delta = 0$ as a mere expedient to exploit the peculiar property of the statistical estimator, shown in Figure 2d: minimising the uncertainty.

However, Figure 2e reveals differences that must be taken into account, being the underlying assumptions different for the two cases. In fact, if a nonzero systematic effect is known, though imprecisely, to exist on line "corrected" of Figure 2e, an estimated non-zero correction is applied, while on the line "uncorrected" is ignored, i.e. set to zero, but keeping the same uncertainty of case (d). Clearly in case (e) a full equivalence of the two uncertainty intervals cannot be guaranteed anymore. Note that this would also be the situation, foreseen in the GUM, when one is not aware of the occurrence of a systematic effect. This circumstance, though obviously not desirable, is not at all rare, and is the main reason of discrepant results: an evidence provided by a large proportion of the inter-comparison of results in either the metrological and testing fields[3].

Some corrections are undisputable: a simple example is the correction for room temperature deviation with respect to a reference one – correction that might have a value zero if the room is set at the nominal reference temperature.

Concerning others, not all experimentalists or statisticians may want to agree about applying a correction, leaving instead the measured value unaltered and enlarging the uncertainty. In this situation, the decision (i) to assign a zero value to the correction or, (ii) to ignore it, may look eventually equivalent, and simply moving the dilemma to the 'enlargement' of uncertainty.

4 Why a distinction between "corrections" and "input quantities"

GUM's recipe is the same of Gauß: should one recognise a systematic effect, a correction must be applied to measured data. However, correcting measured data may clearly be a far from trivial operation.

In fact, the very distinction between quantities relevant to a measurand and corrections to these quantities seems artificial, though customary: the fact that a "correction" may be ignored, or that its value can be zero, has possibly promoted this distinction.

In the author's opinion, there is no actual reasons for such a distinction: by avoiding it, some of the problems that affect the application of "corrections" might be resolved in a simpler and more consistent way. This opinion is substantiated in admissions that can be found in some standards: IEC 60359:2001 [7] "*NOTE 3 – It is understood that the result of the measurement is the best estimate of the value of the measurand, and that all components of uncertainty, including those arising from systematic effects, such as components associated with corrections and reference standards, contribute to the dispersion*" (3.1.4).

[3] This evidence justifies the wide literature on the treatment of "outliers".

However, in the GUM, having admitted in clause (4.1.2) (see Sect. 1 above) that corrections are included in the model, then they are treated differently from the other 'input quantities' when clause (3.2.4) requires that the measured values must preliminarly be "corrected"[4].

To show a contradiction in this issue, let us take a simple explicit model $Y = X_1 \cdot X_2$, where Y is measured indirectly by measuring X_1 and to X_2, with corrections included according to clause (4.1.2):

$$Y = (X_1 + C_1) \cdot (X_2 \cdot C_2) \tag{1}$$

where X_1 and X_2 are two measured 'input quantities', and C_1 and C_2 are 'corrections', respectively to X_1 and to X_2. The corrections can also be measured quantities, or their values and uncertainties can be obtained from Type B evaluations.

The expression (1) in the GUM is the model, according to clause (4.1.2). However, according to clause (3.2.4), in the GUM it is stated: "*It is assumed that the result of a measurement has been corrected for all recognized significant systematic effects and that every effort has been made to identify such effects*". Therefore, model (1) is only the initial model. In fact, after the required corrections are applied, the model (1) should be rewritten as

$$Y = (X_1^* + C_1^*) \cdot (X_2^* \cdot C_2^*) \tag{2}$$

where all quantities on the right are different from those in (1): $X_1^* = X_1 + E(C_1)$; $X_2^* = X_2 \cdot (1 - E(C_2))$; $C_1^* = C_1 - E(C_1)$; $C_2^* = C_2 - E(C_2)$, where, in general, $E(C_1) \neq 0$ and $E(C_2) \neq 0$[5]. The new quantities C_1^* and C_2^* have zero expectation: $E(C_1^*) = 0$ and $E(C_2^*) = 0$.

Expression (2) is the actual GUM model subjected to the uncertainty analysis. The rational is to ensure that the "*Guide treats uncertainty components arising from random effects and from corrections for systematic effects in exactly the same way in the evaluation of the uncertainty of the result of a measurement.*" (E.3). This is correct only if correction values are exact, i.e., exactly compensating the fixed effects. It is a very strong assumption, not actually verified in many circumstances – namely after having compared data from different sources, by means of comparison exercises, like those in the MRA and in the testing field.

A different viewpoint is stressed here: "corrections", when performed, should be simply seen as compensation of deviations of measured from reference states of specified quantities, due to perturbations caused by other influence quantities. These other quantities are also measured, or the deviations from the reference state are estimated from Type B evaluations, exactly as the previous "input quantities": in fact, they *are* part of the input quantities in all respects. They are variables in the GUM, though "bias" is

[4] The clause (4.2.2) would extend to the corrections the meaning of 'input quantities', but in the GUM this becomes confusing – see text.

[5] Obviously here we assume that a correction is not a fixed value as in the VIM [6].

https://doi.org/10.1051/ijmqe/2012024 Published online by Cambridge University Press

often considered as a fixed value – the most notable case is VIM [6], where bias is defined: "*estimate of a systematic measurement error*" (2.18), the latter being "*component of measurement error that in replicate measurements remains constant or varies in a predictable manner*" (2.17).

The existence of an uncertainty, always associated to a correction, prevents to omit any correction from the model, where it would simply be a variable X_i with generally $E(X_i) \neq 0$ – unless, according to the GUM, one wants to label it as non-significant, so dropping also its contribution to uncertainty.

In conclusion, Y should be computed using the full model including the explicit functional form of all corrections, when Y is computed in closed-form. This is possible in many cases.

However, when the model remains in an implicit form, or when one likes to avoid embarking in often complicated analytical computations of uncertainty propagation (involving convolutions of quantities), a suggestion for a simpler way of computing it can be to apply the Monte Carlo method. In this case, in addition to the pdfs of the original input quantities, one includes, for each "correction" term, its pdf with $E(C_i)$ having the value contingently attributed to "bias" – zero, only when justified – i.e. the one that would derive from the GUM clause (3.2.4). In this way, $E(Y)$ directly also embeds the effect of the current "corrections".

From this viewpoint, one does not need anymore to make a distinction between X_i and C_i.

5 Conclusion

The paper has summarised the main reasons why the most difficult issue, trusting a correction, cannot be solved easily nor universally.

However, when the decision is to apply corrections, an edge-cutting solution is to consider, and treat, all "corrections" as they actually are: a type of "input quantities", to be included in the model (1) and treated as any other input quantity.

In many cases, this procedure is more straightforward, not needing the separate the step of first applying the correction value from a separate and subsequent step of taking into account its uncertainty: a more understandable and a simpler solution[6]. In some cases, it may also overcome the dilemma in Section 3 and the need for those methods of uncertainty-increase evaluation [3].

References

1. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, Guide to the expression of uncertainty in measurement (GUM) – Supplement 1: Propagation of distributions using a Monte Carlo method, 2nd edn. (Bureau International des Poids et Mesures (BIPM), Sèvres, 1995), http://www.bipm.org
2. CIPM, *Mutual Recognition of National Measurement Standards and of Calibration and Measurement Certificates Issued by National Metrology Institutes* (BIPM, Sèvres, 1999), http://www.bipm.org
3. F. Pavese, On the difference of meaning of 'zero correction': zero value versus no correction, and of the associated uncertainties, in *Advanced Mathematical and Computational Tools in Metrology IX*, edited by F. Pavese et al. (World Scientific, Singapore, 2012), Vol. 9, pp. 297–309
4. B. Magnusson, S.L.R. Ellison, Treatment of uncorrected measurement bias in uncertainty estimation for chemical measurements, Anal. Bioanal. Chem. **390**, 201–213 (2008)
5. F. Pavese, Methods for the assessment of correction for chemical-impurity effects and related uncertainty in ITS-90 fixed points, namely of e-H_2, Ne, O_2 and Ar, Metrologia **48**, 268–274 (2011)
6. BIPM, IEC, IFCC, ISO, IUPAC, IUPAP, OIML, *International Vocabulary of Basic and General Terms in Metrology* (BIPM, Sèvres, 2008), http://www.bipm.org
7. International Electrotechnical Commission, IEC 60359 (2001)

[6] In this way, the GUM hypothesis is relaxed that only random effects apply after having performed the corrections, in general an exceedingly strong hypothesis.

https://doi.org/10.1051/ijmqe/2012024 Published online by Cambridge University Press