Comparing statistical methods for the correction of the systematic effects and for the related uncertainty assessment

F Pavese
INRIM, 10135 Torino, Italy
E-mail: f.pavese@inrim.it

Abstract. This paper aims at considering the issues related to obtaining reliable and scientifically sound corrections, reviewing statistical methods proposed for either obtaining on the most firm bases the value of a correction and the uncertainty to be associated to it, or to be used should the correction not be applied for insufficient confidence on its value.

1. Introduction
Since 1993, the expression of uncertainty in the field of metrology has been based on the ISO Guide to the expression of uncertainty in measurement (GUM). According to the GUM (3.2.4), “it is assumed that the results of a measurement have been corrected for all recognised significant systematic effects”. The problem of performing a correction is not trivial at all in many cases, and also raises other commonly asked questions that are not addressed clearly by the GUM. Firstly, should we make corrections when the uncertainty in the correction is large? The GUM indicates that only “significant” effects should be corrected, but what does significant mean in practice? Secondly, are there sensible statistical methods for handling uncorrected results? Thirdly, the GUM only considers symmetric distributions; how should we calculate a correction and uncertainty for effects characterised by asymmetric distributions?

This paper aims at considering the issues related to obtaining reliable and scientifically sound corrections, reviewing statistical methods proposed for either obtaining on the most firm bases the value of a correction and the uncertainty to be associated to it, or to be used should the correction not be applied for insufficient confidence on its value.

2. Current methods used for performing a correction
2.1. Modelling the effects of the influence quantities quantitatively (QCM)
Let us assume that we can model the effects of influence quantities by a linear expression:

\[ \Delta Y_{QCM} = \sum_i (m_i X_i), \quad i = 1, \ldots, I \]  

where \( \Delta Y \) is the overall effect of all \( I \) influence quantities \( X_i \) on the measured value \( Y \), and \( m_i \) are generally called the sensitivity coefficients: when a model for the relation \( Y = f(X_1, \ldots, X_I) \) is available,

1 To whom any correspondence should be addressed.
they can be expressed analytically; when it is not, their values are obtained experimentally with the variational method especially popular in engineering. Should a correction be applied for this effect, it has the form

\[ D Y_{\text{QCM}} \approx - \Delta Y, \quad (2) \]

where \( D Y \) indicates an estimate of \( \Delta Y \), preventing from using the equal symbol. The \( X_i \) are random variables, so both are \( \Delta Y \) and \( D Y \). Often they are considered \( \Delta Y = \mathcal{N}(\Gamma, \sigma_q) \) and \( D Y = \mathcal{N}(G, \sigma_c) \), but here the normality assumption is not considered a necessary condition. The quantities \( \sigma_q \) and \( \sigma_c \) are the variances of the overall effects and of the correction, respectively. For \( \Delta Y \), \( \Gamma = E(\Delta Y) \), while for \( D Y \), \( G \) is the correction value: \( \Gamma \) is often called ‘bias’\(^2\) but the definitions of this word are ambiguous, so will not be used in this paper. Since the correction is an estimate, after the correction, not necessarily the corrected \( Y \) indicated in the following as \( \hat{Y} \) is \( \mathcal{N}(\Gamma, \sigma) \) and \( D Y \) is \( \mathcal{N}(G, \sigma_c) \), where \( \sigma \) is a combination of \( \sigma_q \) and \( \sigma_c \).

This fact means that \( G \) is affected by non-uniqueness, not necessarily being described in probabilistic or statistical terms.

Additionally to this fact, the correction \( G \) is affected by an uncertainty, e.g. expressed as a \( \sigma_c \), as indicated above. In general, from Eq. (1)

\[ u^2(DY) \approx u^2(\Delta Y) = \sum_i [u^2(m_i)X_i^2 + u^2(X_i)m_i^2], \quad (3) \]

where the symbol \( \approx \) is used since the estimation also involves the uncertainty. This is a quantitative correction-method (QCM).

### 2.2. Modelling the effects of the influence quantities qualitatively (NQCM)

In the total absence of detailed information about the effect of each and all the influence factors, one can only resort to a non-quantitative correction-method (NQCM).

Extreme values (large and small) of the possible variability of the influence quantities are preferentially removed, and the experimenter is making a judgment either about the overall result of all possible variations or about the variation of each individual influence quantity. As a result, it is assumed that total effect lies in the interval

\[ |Y_M - Y_m| < \Delta Y_{\text{NQCM}} = m_{\text{ave}}X_{\text{tot}}, \quad (4) \]

where \( Y_M \) and \( Y_m \) are the bounds of the variation interval of \( Y \), max and min respectively, and depending on the circumstances, one may be able to estimate separately the total effect of the influence parameters \( X_{\text{tot}} \) and the average sensitivity coefficient \( m_{\text{ave}} \) of the influence parameters over \( Y \), or can only estimate the product \( m_{\text{ave}}X_{\text{tot}} \). Equation (4) transforms in a correction equation as follows

\[ D Y_{\text{NQCM}} = -m_{\text{ave}}X_{\text{tot}}. \quad (5) \]

If one assumes that the interval \( |Y_M - Y_m| \) is characterised by a rectangular (uniform) probability distribution centred on \( Y \), then the value of the correction to be applied is zero, so \( E(Y) = E(\hat{Y}) \), and the uncertainty is given by

\[ u(\Delta Y_{\text{NQCM}}) = (m_{\text{ave}}X_{\text{tot}})/\sqrt{3}. \quad (6) \]

\(^2\) The symbol \( B \) is not used here because in Greek and Latin alphabets it looks identical.
2.3. Hybrid methods
In most cases, more information is known than implied by the applications of the NQCM method. Ideally, if such knowledge is integrated into the uncertainty assessment, it ought to be practical to achieve lower uncertainties than obtained using the NQCM method.

2.3.1. Combined QCM-NQCM method. If the knowledge on the effect of all the influence quantities is partial, there may be useful data for some of them. It is therefore possible to combine the QCM method for the known effects, leaving the application of the NQCM for the remaining individually unknown effects.

2.3.2. One-sided NQCM method. The problem of assessing uncertainty in the presence of effects of known sign and unknown magnitude frequently occurs. This is a special case where the resulting overall effect of the influence quantities may or may not be symmetric, but is known to not be centred about zero.

Assuming that all effects are such that they influence $Y$ with the same sign, individually or overall, the possible variability of $Y$ is bounded according to the following expression, showing the case where the effect is only negative (PQCM means partially quantitative correction method)

$$0 < (Y_M - Y_m) < \Delta Y_{PQCM} = m_{ave} \cdot X_{tot}.$$  \hspace{1cm} (7)

If one assumes that the interval $(Y_M - Y_m)$ is characterised by a rectangular (uniform) probability distribution centred on $Y$, then the value of the correction to be applied is

$$DY_{PQCM} = -(m_{ave} \cdot X_{tot})/2,$$  \hspace{1cm} (8)

and the uncertainty is

$$u(DY_{PQCM}) = (m_{ave} \cdot X_{tot})/2\sqrt{3},$$  \hspace{1cm} (9)

which yields a non-zero correction and an uncertainty half that of the straight NQCM approach. The uncertainty interval should additionally be subjected to some bounds. In fact, because of the one-sided effects of all relevant impurities, $(Y_M - Y_m)$, could not exceed zero (the reverse in the case of all-positive effects); thus, the uncertainty should not extend outside the interval $(0, u_\pm)$. This may bring to an asymmetric interval.

3. Current methods used when not performing a correction
In this case, the experimenter is choosing to not perform the correction, but the uncertainty is somewhat increased by taking into account the variability of the influence quantities.

3.1. Gauß & Legendre approach
The basic approach is illustrated in [1], said to take as reference the independent work of Gauß and Legendre on this subject matter. The derivation of the uncertainty $u$ is based on the assumption that, when $y'$ is an estimate of $Y$, "the best estimate of a quantity $Y$ is determined by minimising the expression

$$E(Y - y')^2 = \text{Var}(Y) + [E(Y) - y']^2 \quad [1],$$  \hspace{1cm} (10)

deriving directly from the general expression, for a random variable $x$, of the best estimate of the measurand, the expectation $m = E(x) = \int x \cdot p(x) \, dx$, and variance $\text{Var}(x) E(x - m) = \int (x - m)^2 \cdot p(x) \, dx = E(x^2 + m^2 - 2xm) = E(x)^2 - m^2$. Obviously the value of $y'$ minimising (10) would be for $y' = E(Y)$.

When "it might be more convenient to estimate the measurand by a value $y^*$ that is not equal to $y = E(Y)' [1]", the standard uncertainty associated with $y^*$ is thus
\[ u^2 (y^*) = u(y)^2 + (y - y^*)^2 = u(y)^2 + DY^2. \] (11)

DY is, as before, the estimated value of the correction. Then, adapting \[2\] to the notation of this paper, for no correction being performed, if \( M^* \) is “the quantity that represents the uncorrected measurand” and “\( q \) is the best estimate of \( Q \), then obviously the value \( m^* = q \) is the best estimate of \( M^* \) and its associated standard uncertainty becomes \( u_{m^*}(m^*) = u(q) \). But, if \( m^* \) is taken as an estimate of \( M \) (omissis)
\[ u^2_{m^*}(m^*) = u^2(m) + (m - m^*)^2 = u^2(q) + u^2(DY) + DY^2. \] (12)

Note the difference between \( u_{m^*}(m^*) \) and \( u_{m^*}(m^*) \). Equation (12) is the basis for the expression (13) that follows, omitting to express also the observation uncertainty:
\[ u^2_{m^*}(m^*) - u_{m^*}(m^*) = u^2(DY) + DY^2. \] (13)

Therefore, if no correction is applied, its estimated value is still required to compute the resulting increase of uncertainty. This may look to indicate that is useless to avoid the correction: however, the use of an uncertain correction value is much less critical in the computation of uncertainty than for altering the measured value by applying the correction.

3.2. Other approaches
On the other hand, the approach in Section 3.1 \[1\] brings to the paradox that, in the expression of the expanded uncertainty \( U = ku \) (with \( k \approx 2 \)), also the estimated correction \( DY \) is multiplied by \( k \). This occurrence is not considered correct by other authors, who preferred to directly express the expanded uncertainty. Three main variations are available from the literature.

3.2.1. RSSU approach. This approach \[3\] differs from the previous approach, called RSSu, because considers more correct to express directly the expanded uncertainty \( U \), thus considers the value of the correction \( DY \) not included in the standard uncertainty but in the expanded uncertainty, whence having to be divided by the coverage factor \( k \) in the expression of \( u \) in Eq. (13).

3.2.2. SUMU approach. This approach \[4\] first combines in quadrature the uncertainty of the statistical components (the observation and the correction) to obtain the standard uncertainty \( u \), then computes the corresponding expanded uncertainty \( U = ku \). Finally, the value of the correction \( DY \) is added to it: in this way, the correction is treated as weighting \( 1/k \) in \( u \). SUMU also constrains the uncertainty interval, which is asymmetric, to not extend to negative values.

3.2.3. SUMUMAX approach \[4\] is similar to the SUMU approach, except that transforms the interval into a symmetric one by using \( SUMU_\text{MAX} \) value.

4. Further contents of the presentation and final paper
The merits of the above methods, developed in both fields of metrology and testing, will be compared and some conclusions attempted.

5. References
[1] Lira I H, Wöger W 1998 Meas Sci Technol 9 1067–1073
[2] Lira I H, Wöger W 1998 Meas Sci Technol 9 1010–1011
[3] Magnusson B, Ellison S L R 2008 Anal Bioanal Chem 390 201–213
[4] Phillips S D, Eberhardt K R, Parry B 1997 J Res Nat Inst Stand Technol 102 577–585