Unknown systematic errors and the method of least squares

M. Grabe [1]

Abstract. This paper discusses the impact of unknown systematic errors in the method of least squares with reference to the error models laid down in the *GUM* (ISO *Guide*) and an alternative.

The paper shows that the implications of least squares adjustments depend heavily on the error model called upon. Consequently, it is worthwhile scrutinizing these dependencies aiming at the role of the Gauss-Markoff theorem, the estimation of measurement uncertainties, and, as the alternative error model is based on true values, at the ideas of traceability and key comparison reference values.

1. Introduction

In contrast to the *GUM*, the error model which will be discussed here, and which will simply be called *the alternative error model*, assigns true values to the physical quantities to be measured. This approach may be seen to be backed by the idea that physical laws in general as well as relations between physical constants in particular are to be expressed in true values. In the same way, metrology should estimate true values even if the definition of the latter remains abstract. Strictly speaking, true values scarcely exist in reality, as measurands are subjected to microscopic fluctuations. We may however presuppose in most cases that the latter remain insignificant compared with the accuracy of measurement actually attainable. Should this prerequisite not apply, additional systematic errors would have to be assigned to the measurands. But this normally is expected to happen only rarely.

By their very nature, stationary measurement processes separate the flow of random and systematic errors. So the repeated measurements of a stationary experimental set up should be modelled by an error equation of the form

$$\begin{cases} x_l = x_0 + \varepsilon_l + f; \quad l = 1,...,n \\ -f_s \le f \le f_s \end{cases}. \tag{1.1}$$

[1] The contribution expresses the personal opinion of the author. Please refer to Dr. Michael Grabe, Am Hasselteich 5, 38104 Braunschweig, Germany, michael.grabe@ptb.de

The quantities $\varepsilon_l;\ l=1,...,n$ denote random errors and the quantity f designates a systematic error unknown as regards magnitude and sign. All we know is that f lies somewhere between its bounds $\pm f_s$. The unknown true value x_0 is an abstraction. Accepting the latter, x_0 would assume the role of a reference point for all subsequent considerations. Otherwise the basic idea of traceability would be lost and with that the possibilities of properly using and properly testing physical laws.

From the error model (1.1) we deduce: The sequence of random errors scatters with respect to the expectation

$$\mu_x = E\{X_l\} = x_0 + f\ , \tag{1.2}$$

where, formally, a random variable X_l has been assigned to the sequence x_l, $l=1,...,n$. For the sake of simplicity, the curly brackets substitute the probability integral, see e.g. [6]. Herewith, (1.1) transforms into the identity,

$$x_l = x_0 + (x_l - \mu_x) + f\ ; \quad l=1,...,n\ . \tag{1.3}$$

Summing up over $l=1,...,n$ yields

$$\bar{x} = x_0 + (\bar{x} - \mu_x) + f\ . \tag{1.4}$$

In the following, we will investigate the consequences of introducing true values and time-constant (unknown) systematic errors in least squares adjustments. As is well known, the *GUM* recommends treating systematic errors in probabilistic terms by means of postulated rectangular distribution densities. In contrast to this, the alternative error model considers systematic errors to cause biases and aims at bringing their influence to bear by means of *worst-case* estimations.

2. The method of least squares

The method of least squares can be introduced via the idea of orthogonal projection. Let the real numbers $a_{ik};\ i=1,...,m;\ k=1,...,r$ define an $m\times r$ matrix $\boldsymbol{A}$ of rank r. The r unknowns $\beta_k;\ k=1,...,r$ to be estimated by least squares will be condensed within an $r\times 1$ column vector $\boldsymbol{\beta}$. Finally, let the input data $x_i;\ i=1,...,m$ be the elements of an $m\times 1$ column vector $\boldsymbol{x}$ of observations. As the latter are erroneous, only an approximate solution to the linear system

$$\boldsymbol{A\beta} \approx \boldsymbol{x} \tag{2.1}$$

exists. We require that this system should be readable in terms of true values, i.e. (2.1) should define the overdetermined and *consistent* system

$$\boldsymbol{A\beta}_0 = \boldsymbol{x}_0\ , \tag{2.2}$$

where

$$\boldsymbol{x}_0 = \left(x_{0,1} \quad ... \quad x_{0,m}\right)^T\ , \tag{2.3}$$

designates the vector of true input data and

(2.4) $$\boldsymbol{\beta}_0 = \left(\beta_{0,1} \quad \ldots \quad \beta_{0,r}\right)^T$$

the true solution vector. The symbol T denotes the transpose of any vector or matrix. Evidently, should (2.2) not exist, the adjustment would be based on an ill-defined physical problem. We clearly see that physics requires us to refer to true values. Naturally, metrology has to do the same. In fact, taking reference to true values expresses the basic idea of traceability.

As an illustrative example, let us calculate the *grand mean* of a given set of means

(2.5) $$\bar{x}_i = \frac{1}{n}\sum_{l=1}^{n} x_{il}\ ; \quad i = 1,...,m\ .$$

For each of the means we assume the *same number* of repeated measurements. As may be shown, this measure is supported by statistics, viz. by the joint distribution of the empirical moments of second order.

Forming the *grand mean*, systems (2.1) and (2.2) change into

$$\begin{pmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{pmatrix} \beta \approx \begin{pmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \ldots \\ \bar{x}_m \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 1 \\ 1 \\ \ldots \\ 1 \end{pmatrix} \beta_0 = \begin{pmatrix} x_{0,1} \\ x_{0,2} \\ \ldots \\ x_{0,m} \end{pmatrix}.$$

The right hand system reveals that averaging of a set of means is admissible *if and only if* their respective true values $x_{0,i};\ i = 1,...,m$ are *identical* as the system would otherwise neither mathematically nor physically be defined. To assume identically true values is an indispensable prerequisite for defining so-called *Key Comparison Reference Values*.

On the other hand, a *grand mean* (or a mean of means) could be defined *ad hoc*, i.e. without making reference to a self-consistent system of the kind quoted in (2.2). Let us consider what happens if this is done: Given two masses of $m_1 = 1/4\,kg$ and $m_2 = 3/4\,kg$, respectively, to be accurate to about $1\,mg$. Their mean amounts to $\bar{m} = 1/2\,kg$. The associated uncertainty exceeds the uncertainties of its components by the factor 250,000. The reason for this huge discrepancy is obvious: The uncertainty of the mean has been caused by the difference between the two true values. We clearly see: As long as the true values of the components of a mean of means do not coincide, the associated uncertainty is inflated by the differences between the true values in question. Compared with this, measurement uncertainties should be based on *measurement errors* and not on differences between true values.

After all, according to the alternative formalism, the *grand mean* is given by

(2.6) $$\bar{\beta} = \sum_{i=1}^{m} w_i \bar{x}_i\ .$$

Here, the coefficients w_i denote weighting factors. The associated uncertainty is given by

(2.7) $$u_{\bar{\beta}} = \frac{t_P(n-1)}{\sqrt{n}} \sqrt{\sum_{i,j}^{m} w_i w_j s_{ij}} + \sum_{i=1}^{m} w_i f_{s,i} \,.$$

The factor t_P refers to Student's statistics while the quantities

(2.8) $$s_{ij} = \frac{1}{n-1} \sum_{l=1}^{n} (x_{il} - \bar{x}_i)(x_{jl} - \bar{x}_j); \quad i,j = 1,...,m$$

designate the elements of the empirical variance-covariance matrix of the input data. One might suspect the second term on the right hand side of (2.7) overestimates the contribution due to systematic errors, much the more so as this term seems to increase with the number of means. However, on account of the weighting factors, this term remains of a reasonable order of magnitude, i.e. there is no overestimaton, irrespective of how many means have been averaged. This may easily be demonstrated by examples.

Let us return to the least squares solution of the overdetermined, inconsistent system (2.1). Let the input data of the adjustment be arithmetic means

(2.9) $$\bar{x}_i = x_{0,i} + (\bar{x}_i - \mu_i) + f_i; \quad i = 1,...,m \,.$$

For convenience, we assign a column vector $\bar{\boldsymbol{x}} = (\bar{x}_1 \quad \bar{x}_2 \quad ... \quad \bar{x}_m)^T$ to $\bar{x}_i; \, i = 1,...,m$, so that

(2.10) $$\bar{\boldsymbol{x}} = \boldsymbol{x}_0 + (\bar{\boldsymbol{x}} - \boldsymbol{\mu}) + \boldsymbol{f} \,.$$

Then (2.1) reads

(2.11) $$\boldsymbol{A\beta} \approx \bar{\boldsymbol{x}} \,.$$

As is well known, the method of least squares substitutes a vector $\boldsymbol{P}\bar{\boldsymbol{x}}$ for $\bar{\boldsymbol{x}}$, where $\boldsymbol{P}\bar{\boldsymbol{x}}$ is the orthogonal projection of $\bar{\boldsymbol{x}}$ onto the column space of $\boldsymbol{A}$ so that

(2.12 $$\boldsymbol{A}\bar{\boldsymbol{\beta}} = \boldsymbol{P}\bar{\boldsymbol{x}} \,.$$

The implied projection operator is given by $\boldsymbol{P} = \boldsymbol{A}(\boldsymbol{A}^T\boldsymbol{A})^{-1}\boldsymbol{A}^T$, see e.g. [5]. Clearly, (2.12) yields the least squares solution vector

(2.13) $$\bar{\boldsymbol{\beta}} = \left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}\boldsymbol{A}^T\bar{\boldsymbol{x}} = \boldsymbol{B}^T\bar{\boldsymbol{x}} \,,$$

setting $\boldsymbol{B} = \boldsymbol{A}\left(\boldsymbol{A}^T\boldsymbol{A}\right)^{-1}$. Furthermore, we define $\bar{\boldsymbol{\beta}} = \left(\bar{\beta}_1 \quad ... \quad \bar{\beta}_r\right)^T$.

In every respect, the orthogonal projection *is* the method of least squares. Evidently, this projection is independent of the error model assigned to the input data $\bar{x}_i$, $i = 1,...,m$, i.e. the projection yields a solution vector for *any set* of input data (be it convenient or absurd) as long as the matrix $\boldsymbol{A}$ is of rank r.

Incidentally, from (2.2) and (2.13) we derive the true solution vector

(2.14) $$\boldsymbol{\beta}_0 = \boldsymbol{B}^T\boldsymbol{x}_0 \,.$$

In what follows, the uncertainties $u_{\bar{\beta}_k}$ of the components $\bar{\beta}_k$; $k = 1,...,r$ of the solution vector $\bar{\boldsymbol{\beta}}$, as defined in (2.13), will always be specified with reference to the true values $\beta_{0,k}$.

3. The role of the Gauß-Markoff theorem

The Gauß-Markoff theorem reads: Among the class of estimators being linear in the input data, those designed according to least squares yield minimum diagonal elements of the variance-covariance matrix of the solution vector.

To the author's knowledge, the *GUM* does not presuppose the validity of the Gauß-Markoff theorem. Nevertheless, least squares adjustments carried out in accordance with the *GUM* at least implicitly refer to the results of this theorem. As the *GUM* treats random and systematic errors on essentially the same footing, adjusted systems should imply minimal measurement uncertainties. At the same time, the minimized sums of squared residuals should approximate the number of degrees of freedom of the linear systems considered.

Until so-called unknown systematic errors became a subject of broad and intense discussion, say about 1978, [1], true values were unanimously understood as statistical expectations of the kind

$$\mu_k = E\{\overline{\beta}_k\}; \quad k = 1,...,r\,. \tag{3.1}$$

In contrast to this, the alternative error model introduces biased estimators

$$\mu_k = E\{\overline{\beta}_k\} = \beta_{0,k} + bias(\overline{\beta}_k); \quad k = 1,...,r\,, \tag{3.2}$$

where $bias(\overline{\beta}_k)$ designates time-constant, non-statistical quantities caused by unknown systematic errors. Indeed, these different views reflect the essence of the impact of unknown systematic errors in least squares. As is known, any bias not only invalidates the Gauss-Markoff theorem but also abolishes the classical test for consistency of the input data. If $\overline{Q}_{\min}$ denotes the minimized sum of squared residuals, we should have

$$E\{\overline{Q}_{\min}\} = m - r\,, \tag{3.3}$$

as the *Guide's* basic assumption is to treat both random and unknown systematic errors on a probabilistic basis. As is well known, empirical findings as presented e.g. in [7] hardly confirm statement (3.3). In accord with this observation, the alternative error model yields something very different. At the moment, let it suffice just to note

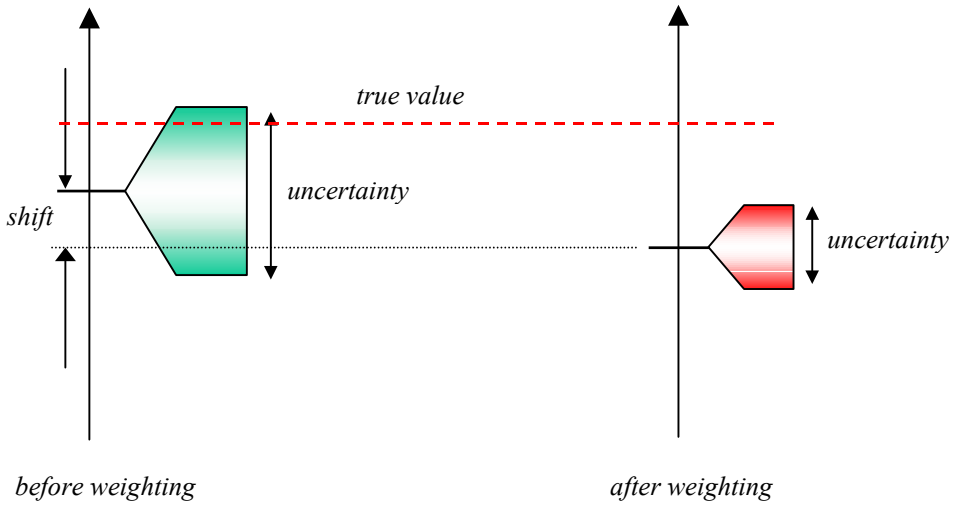
$$E\{\overline{Q}_{\min}\} \neq m - r\,. \tag{3.4}$$

For practical reasons, the details of this expression will be given at the end of the section. Evidently, the breakdown of the Gauß-Markoff theorem invalidates not only the conventional consistency test (3.3) but also the conventionally assumed *optimal properties* of least squares estimators. After all, we have to look for an alternative procedure, of how to weight the linear system in order to extract as much information from the input data as possible. Let us recall that the consequences of weighting procedures are twofold: They shift the numerical values of the estimators and they reduce the uncertainties of the latter. Nevertheless, the complexes

$$(estimator \pm measurement\ \ uncertainty)_k\ ;\ \ k = 1,...,r \tag{3.5}$$

should always localize the true values of the estimators. While adjustments carried out on the basis of the alternative error model strictly fulfil this requirement, the *GUM* leaves open what

happens.



The alternative error model *ad hoc* introduces a weighting matrix $\boldsymbol{G}$, preferably in the form of a (non-singular) diagonal matrix

$$\boldsymbol{G} = diag\{g_1, \quad g_2, \quad \ldots \quad g_m\}. \tag{3.6}$$

Multiplying (2.11) and (2.2) from the left by $\boldsymbol{G}$, we find $\boldsymbol{GA\beta} \approx \boldsymbol{G\bar{x}}$ and $\boldsymbol{GA\beta}_0 = \boldsymbol{Gx}_0$, respectively, so that

$$\begin{cases} \bar{\boldsymbol{\beta}} = \tilde{\boldsymbol{B}}^T \boldsymbol{G\bar{x}} \\ \boldsymbol{\beta}_0 = \tilde{\boldsymbol{B}}^T \boldsymbol{Gx}_0 \end{cases}; \qquad \tilde{\boldsymbol{B}} = \boldsymbol{GA}\left[\left(\boldsymbol{GA}^T\right)\left(\boldsymbol{GA}\right)\right]^{-1}. \tag{3.7}$$

From this we deduce that weighting procedures do not shift the true solution vector $\boldsymbol{\beta}_0$. Consequently, the alternative error model always yields uncertainty intervals of the kind

$$\bar{\beta}_k - u_{\bar{\beta}_k} \le \beta_{0,k} \le \bar{\beta}_k + u_{\bar{\beta}_k}, \qquad k = 1,\ldots,r; \tag{3.8}$$

the $u_{\bar{\beta}_k}$ will be specified in section 4. As has been outlined above, weighting procedures shrink uncertainties and shift estimators. Thus, it is of utmost importance to know whether the complexes quoted as symbols in (3.5) and mathematically written in (3.8), confirm the localization of the true values $\beta_{0,k}$; $k = 1,\ldots,r$.

After all, the weightings of the alternative error model fulfil this requirement irrespective of which (non-singular) matrix of weights has been chosen (more or less advantageous). We are therefore in a position to vary the elements g_i of the matrix $\boldsymbol{G}$ by *trial and error* to minimize the uncertainties $u_{\bar{\beta}_k}$ of the adjusted least squares estimators. In a certain sense, this kind of this optimization may be seen to replace the Gauss-Markoff theorem we have abandoned on account of our intention, to understand unknown systematic errors as time-constant quantities.

The optimization procedure is not necessarily critical. To start, we simply set

$$g_i = 1/u_{\bar{x}_i}; \qquad i = 1,\ldots,m,$$

where the $u_{\bar{x}_i}$ denote the uncertainties of the input data $\bar{x}_i$. A practical example will be pre-

sented in section 5.

Finally, to complete the picture, we insert the missing details of expression (3.4). Given there are no dependencies between the components of $\bar{\boldsymbol{x}}$, we have

$$E\{\overline{Q}_{\min}\} = \sum_{i=1}^{m} g_i^2 \frac{\sigma_i^2}{n} - \sum_{i=1}^{m} g_i^2 \frac{\sigma_i^2}{n} \tilde{p}_{ii} + \left[(\boldsymbol{Gf})^T (\boldsymbol{Gf}) - (\boldsymbol{Gf})^T \tilde{\boldsymbol{P}} (\boldsymbol{Gf}) \right]. \tag{3.9}$$

The $\tilde{p}_{ii}$ denote the diagonal elements of the projection operator of the *weighted* adjustment,

$$\tilde{\boldsymbol{P}} = (\boldsymbol{GA}) \left[(\boldsymbol{GA})^T (\boldsymbol{GA}) \right]^{-1} (\boldsymbol{GA})^T ,$$

and the σ_i the theoretical variances of the input data.

4. The alternative uncertainties

Inserting (2.10) into (3.7) we arrive at

$$\bar{\boldsymbol{\beta}} = \tilde{\boldsymbol{B}}^T \boldsymbol{G} \left[\boldsymbol{x}_0 + (\bar{\boldsymbol{x}} - \boldsymbol{\mu}) + \boldsymbol{f} \right]. \tag{4.1}$$

On the right hand side of

$$\bar{\boldsymbol{\beta}} = \boldsymbol{\beta}_0 + \tilde{\boldsymbol{B}}^T \boldsymbol{G} (\bar{\boldsymbol{x}} - \boldsymbol{\mu}) + \tilde{\boldsymbol{B}}^T \boldsymbol{Gf} , \tag{4.2}$$

the second term is due to random errors while the third term

$$\tilde{\boldsymbol{f}}_{\bar{\beta}} = \tilde{\boldsymbol{B}}^T \boldsymbol{Gf} \tag{4.3}$$

expresses the influence of systematic errors. This term also appears in the expectation

$$\boldsymbol{\mu}_{\bar{\beta}} = E\{\bar{\boldsymbol{\beta}}\} = \boldsymbol{\beta}_0 + \tilde{\boldsymbol{B}}^T \boldsymbol{Gf} . \tag{4.4}$$

Let the $\tilde{b}_{ik}$; $i = 1,...,m$; $k = 1,...,r$ denote the elements of the $m \times r$ matrix $\tilde{\boldsymbol{B}}$. Then, the components

$$\tilde{f}_{\bar{\beta}_k} = \sum_{i=1}^{m} \tilde{b}_{ik} g_i f_i ; \quad k = 1,...,r \tag{4.5}$$

lead us to the worst-case estimations

$$\tilde{f}_{s,\bar{\beta}_k} = \sum_{i=1}^{m} \left| \tilde{b}_{ik} \right| g_i f_{i,s} ; \quad k = 1,...,r . \tag{4.6}$$

We assess the empirical variance-covariance of the solution vector $\bar{\boldsymbol{\beta}}$ in analogy to (2.8), i.e. we resolve its components

$$\bar{\beta}_k = \sum_{i=1}^{m} \tilde{b}_{ik} g_i \bar{x}_i = \frac{1}{n} \sum_{l=1}^{n} \left[\sum_{i=1}^{m} \tilde{b}_{ik} g_i x_{il} \right] = \frac{1}{n} \sum_{l=1}^{n} \bar{\beta}_{kl} ; \quad k = 1,...,r \tag{4.7}$$

to find the quantities

$$\bar{\beta}_{kl} = \sum_{i=1}^{m} \tilde{b}_{ik} g_i x_{il} ; \quad l = 1,...,n . \tag{4.8}$$

The differences

$$\overline{\beta}_{kl} - \overline{\beta}_k = \sum_{i=1}^{m} \widetilde{b}_{ik} g_i \left(x_{il} - \overline{x}_i \right); \quad k = 1,...,r; \; l = 1,...,n$$

lead us to the elements

$$s_{\overline{\beta}_k \overline{\beta}_{k'}} = \frac{1}{n-1} \sum_{l=1}^{n} \left(\overline{\beta}_{kl} - \overline{\beta}_k \right) \left(\overline{\beta}_{k'l} - \overline{\beta}_{k'} \right); \quad k, k' = 1,...,r \tag{4.9}$$

of the empirical variance-covariance matrix $\boldsymbol{s}_{\overline{\boldsymbol{\beta}}}$ of the solution vector $\overline{\boldsymbol{\beta}}$. We have

$$s_{\overline{\beta}_k \overline{\beta}_{k'}} = \frac{1}{n-1} \sum_{l=1}^{n} \left[\sum_{i=1}^{m} \widetilde{b}_{ik} g_i \left(x_{il} - \overline{x}_i \right) \right] \left[\sum_{j=1}^{m} \widetilde{b}_{jk'} g_j \left(x_{jl} - \overline{x}_j \right) \right] = \sum_{i,j}^{m} \widetilde{b}_{ik} \widetilde{b}_{jk'} g_i g_j s_{ij}, \qquad s_{\overline{\beta}_k \overline{\beta}_k} \equiv s^2_{\overline{\beta}_k};$$

in matrix notation

$$\boldsymbol{s}_{\overline{\boldsymbol{\beta}}} = \widetilde{\boldsymbol{B}}^T \left(\boldsymbol{G} \boldsymbol{s} \boldsymbol{G}^T \right) \widetilde{\boldsymbol{B}}. \tag{4.10}$$

Evidently, the quantities β_{kl} entering (4.7) are independent, irrespective of whether there are dependencies between the means $\overline{x}_i$, $i = 1,...,m$, simply because we do not admit dependencies between repeated measurements of one and the same quantity.

Just as for a normally distributed random variable X, given its realizations are independent, the quotient

$$\frac{\overline{X} - \mu_x}{S / \sqrt{n}}; \qquad E\{\overline{X}\} = \mu_x$$

represents a Student's t, the same applies to

$$\frac{\overline{\beta}_k - \mu_{\overline{\beta}_k}}{S_{\overline{\beta}_k} / \sqrt{n}}; \qquad E\{\overline{\beta}_k\} = \mu_{\overline{\beta}_k}. \tag{4.11}$$

After all, the intervals

$$\overline{\beta}_k - t_P(n-1) \frac{s_{\overline{\beta}_k}}{\sqrt{n}} \le \mu_{\overline{\beta}_k} \le \overline{\beta}_k + t_P(n-1) \frac{s_{\overline{\beta}_k}}{\sqrt{n}}; \quad k = 1,...,r, \tag{4.12}$$

where $t_P(n-1))$ denotes the Student factor, localize the expectations $\mu_{\overline{\beta}_k}$ of the $\overline{\beta}_k$ with probability P.

As long as the means $\overline{x}_i$, $i = 1,...,m$ are mutually independent, it remains immaterial which of the repeated measurements

$$x_{1,l_1}, x_{2,l_2}, ..., x_{m,l_m}$$

out of the sequences

$$\underbrace{x_{1,1}, x_{1,2}, ..., x_{1,n}}_{\overline{x}_1}; \qquad \underbrace{x_{2,1}, x_{2,2}, ..., x_{2,n}}_{\overline{x}_2}; \qquad \underbrace{x_{m,1}, x_{m,2}, ..., x_{m,n}}_{\overline{x}_m}$$

enter β_{kl}; $l = 1,...,n$. The associated redundancy is included in Student's confidence intervals, which, by definition, vary according to the actual samples. I.e. statistically "breathing" confidence intervals do not affect the underlying probability statement.

Should, say $\bar{x}_i$ and $\bar{x}_{i'}$, be dependent, the ordered sequence x_{il}, $x_{i'l}$ may however not be altered. Nevertheless, as has already been stated, even then the β_{kl}; $l = 1,...,n$ are independent.

The linear combination of (4.6) and (4.12) yields the overall uncertainties

$$u_{\bar{\beta}_k} = \frac{t_P(n-1)}{\sqrt{n}} \sqrt{\sum_{i,j}^{m} \tilde{b}_{ik}\tilde{b}_{jk} g_i g_j s_{ij}} + \sum_{i=1}^{m} \left|\tilde{b}_{ik}\right| g_i f_{s,i}; \quad k = 1,...,r \tag{4.13}$$

of the alternative error model. As has been pointed out, they are expected to localize the true values $\beta_{0,k}$; $k = 1,...,r$ nearly with certainty. Again, the worst-case estimations of the systematic errors should scarcely express overestimations as the coefficients $\tilde{b}_{ik}$ act in a manner similar to that of "weight factors" (though differently assembled). After all, both items of (4.13) represent virtually safe estimates.

5. A simulated adjustment

To illustrate the proposed weighting procedure, we consider the following linear system,

$$\begin{array}{rrrcl}
2\beta_1 & +\beta_2 & +\beta_3 & = & \bar{x}_1 \\
-\beta_1 & & +2\beta_3 & = & \bar{x}_2 \\
\beta_1 & +2\beta_2 & +2\beta_3 & = & \bar{x}_3 \\
-3\beta_1 & -\beta_2 & +2\beta_3 & = & \bar{x}_4 \\
\beta_1 & +\beta_2 & +\beta_3 & = & \bar{x}_5
\end{array}$$

Selecting

$$\boldsymbol{x}_0 = (7 \quad 5 \quad 11 \quad 1 \quad 6)^T,$$

we have

$$\boldsymbol{\beta}_0 = (1 \quad 2 \quad 3)^T.$$

To vectors $\boldsymbol{f} = (f_1 \quad f_2 \quad \ldots \quad f_m)^T$ and $\boldsymbol{f}_s = (f_{s,1} \quad f_{s,2} \quad \ldots \quad f_{s,m})^T$ as defined in (2.10) we assign

$$f = (0.07 \quad 0.0035 \quad -0.007 \quad 0.0035 \quad 0.006)$$

and

$$f_s = (0.07 \quad 0.0035 \quad 0.007 \quad 0.0035 \quad 0.006),$$

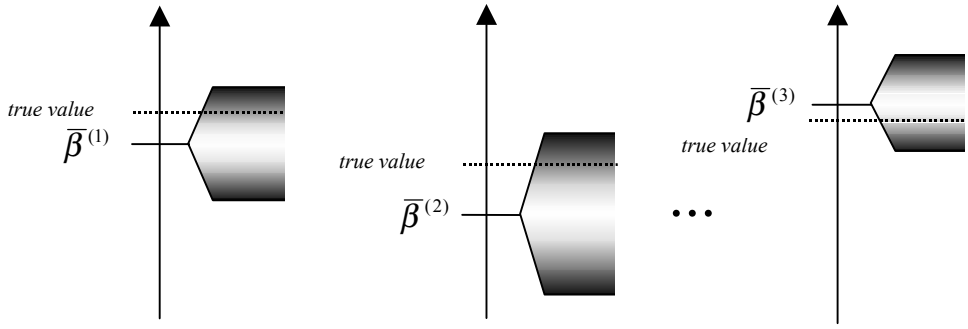
respectively. We assume $n = 10$ repeated measurements for each of the input data $\bar{x}_i$, $i = 1,...,5$ and, for convenience, adjust the associated theoretical variance of the simulated random errors to be equal to the theoretical variance of the systematic error (in the sense of the *Guide's* notion). The diagonal elements of the *Guide's* weighting matrix have been set equal to the reciprocals of the *Guide's* uncertainties of the input data.

	ISO unweighted	ISO weighted	alternative unweighted	alternative weighted
$\overline{\beta}_1 \pm u_{\overline{\beta}_1}$	1.050 ± 0.052	1.012 ± 0.010	1.05 ± 0.08	1.012 ± 0.013
$\overline{\beta}_2 \pm u_{\overline{\beta}_2}$	1.93 ± 0.08	1.979 ± 0.016	1.93 ± 0.11	1.979 ± 0.022
$\overline{\beta}_3 \pm u_{\overline{\beta}_3}$	3.041 ± 0.038	3.010 ± 0.008	3.04 ± 0.06	3.009 ± 0.010

For the *Guide*, 2σ uncertainties have been quoted. Remarkably enough, in the case of weighting, the uncertainties of both error models are scanty.

6. Key Comparisons

Let us consider a round robin to test the consistency of a set of *National Standards* $\overline{\beta}^{(1)}$, $\overline{\beta}^{(2)}$,..., $\overline{\beta}^{(m)}$. Surely, metrology freely admits their true values to deviate from another.



Neither the true values of standards of the same hierarchical level must coincide, nor must their uncertainties overlap – however, the uncertainties must localize the respective true values.

However, the associated uncertainties must localize the respective true values. Each participant calibrates one and the same *travelling standard* T. Providing the true value of the latter remained time-constant, the uncertainties of the calibrations should mutually overlap. A horizontal line, intersecting each of the uncertainty bounds, would indicate compatible calibrations and thus compatible *National Standards* $\overline{\beta}^{(1)}$, $\overline{\beta}^{(2)}$,..., $\overline{\beta}^{(m)}$, though their true values may deviate from one another in any size whatever.

However, key comparisons do more. One option is to define a weighted *grand mean*,

$$\overline{\beta} = \sum_{i=1}^{m} w_i \overline{T}^{(i)}, \tag{6.1}$$

called the *Key Comparison Reference Value*, and to scrutinize the uncertainties of the m differences

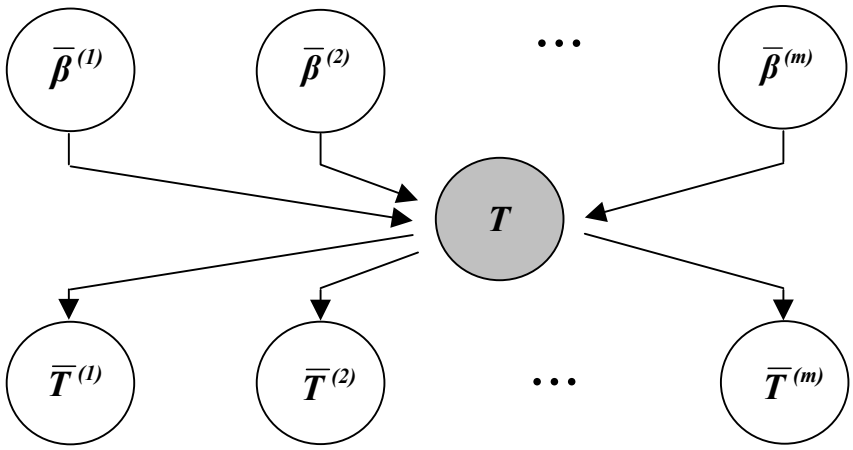
$$\overline{d}_i = \overline{T}^{(i)} - \overline{\beta}\,; \quad i = 1,...,m\,. \tag{6.2}$$

The alternative error model yields

$$\text{(6.3)}\qquad \begin{cases} \left|\overline{T}^{(i)} - \overline{\beta}\right| \le u_{\overline{d}_i} \\ u_{\overline{d}_i} = \dfrac{t_P(n-1)}{\sqrt{n}} \sqrt{s_i^2 - 2\sum_{j=1}^{m} w_j s_{ij} + \boldsymbol{w}^T \boldsymbol{s}\, \boldsymbol{w} + f_{s,i} - 2 w_i f_{s,i} + \sum_{j=1}^{m} w_j f_{s,j}} \end{cases}.$$

Formally reducing the assumptions of the alternative error model to that of the *GUM* turns (6.3) into

$$u^2(d_i) = u^2\left(T^{(i)}\right) - u^2(\overline{\beta}).$$



Round robin of a set of m standards $\overline{\beta}^{(i)}$; $i = 1,...,m$ realized by a travelling standard T. The calibrations $\overline{T}^{(i)}$ should be consistent .

The differences $\overline{d}_i$ and their uncertainties $u_{\overline{d}_i}$ are not easy to interpret. With respect to the *GUM* it might happen that some of the calibrations localize the true value of the travelling standard but others not. As each of the calibrations enters the *Key Comparison Reference Value* $\overline{\beta}$, it might happen that incorrect calibrations dominate the correct ones, so that correct calibrations may appear incorrect while erroneous ones may appear correct.

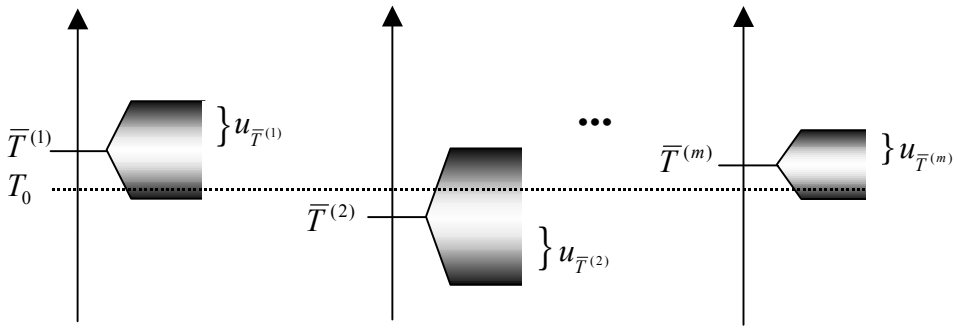
Consequently, with respect to *GUM*'s philosophy not to localize true values, it appears questionable to test $\overline{T}^{(i)}$ against $\overline{\beta}$, when $\overline{T}^{(i)}$ contributes to $\overline{\beta}$.

7. Conclusion

In his introductory words to the panel discussion "Should least squares adjustments of the fundamental constants be abolished?" [9], P.L. Bender addressed one of the arguably most delicate problems of metrology, namely what has to be done if the uncertainties of physical quantities, accurately estimated, clearly disagree.

It goes without saying that the measure of discrepancy depends on the measure of uncertainty referred to: Are we looking for 1σ, 2σ or $k_p\sigma$ uncertainties? As long as the extensions of the uncertainty intervals are not backed up by a strictly accountable concept, there will be no

objective measure of discrepancy. In the opinion of the author, such a measure will not be within reach until the concept of *true values* has been (re)introduced, until random and systematic errors will be kept *strictly separate* and, finally, until uncertainties will be defined on a premise in the sense:



Round Robin: In case of consistency of the *National Standards* $\overline{\beta}^{(i)}$ and the calibrations $\overline{T}^{(i)}$, a horizontal line would intersect each of the uncertainties $u_{\overline{T}^{(i)}}$.

The uncertainty of a measured quantity defines the smallest interval localizing the true value of the physical quantity in question (nearly) with certainty.

We conclude that the problems revealed by the aforementioned panel discussion might have arisen from the error model in use at the time. The *Guide* has not altered the essence of that notion, although the commonly employed procedures of data evaluation and adjustment, which, after all, were originally developed by K.F. Gauß and his contemporaries, seem to suffer from the progress of measurement techniques, requiring the explicit definition of true values and the general inclusion of biases.

References

[1] *Über die Fortpflanzung zufälliger und systematischer Fehler*, Seminar on the statement of measurement uncertainty, February 20 and 21, 1978, PTB Braunschweig.

[2] ISO, *Guide to the Expression of Uncertainty in Measurement*, 1993. 1, Rue de Varambé, P.O. Box 56, CH-1211 Geneva 20, Switzerland.

[3] Grabe, M., *Principles of "Metrological Statistics,"* Metrologia **23** (1986/87) 213-219.

[4] Grabe, M., *Estimation of measurement uncertainties – an alternative to the* ISO *Guide*, Metrologia **38** (2001) 97-106.

[5] Seber, G.A.F., *Linear Regression Analysis*, John Wiley & Sons, 1977.

[6] Papoulis, A. *Probability, Random Variables and Stochastic Processes*, McGraw-Hill, 1965.

[7] Cohen, E.R. and Taylor, B.R.: *The 1986 Adjustment of the Fundamental Physical Constants*, CODATA BULLETIN No. 63 (1986).

[8] Cox, Maurice,G.: *The evaluation of key comparison data,* Metrologia, **39**(2002)589-595

[9] Bender, P.L. et al. *Should Least Squares Adjustment of the Fundamental Constants be Abolished?* In *Precision Measurement and Calibration*, NBS Special Publication 343, Washington, D.C., 1971.