Measurement 187 (2022) 110340

Contents lists available at ScienceDirect

Measurement

journal homepage: www.elsevier.com/locate/measurement

The GUM perspective on straight-line errors-in-variables regression

Katy Klauenberg [a,*], Steffen Martens [b], Alen Bošnjaković [c], Maurice G. Cox [d], Adriaan M.H. van der Veen [e], Clemens Elster [a]

[a] Physikalisch-Technische Bundesanstalt, Abbestrasse 2-12, 10587 Berlin, Germany
[b] Independent researcher, Berlin, Germany
[c] Institute of Metrology of Bosnia and Herzegovina, Augusta Brauna 2, 71 000 Sarajevo, Bosnia and Herzegovina
[d] National Physical Laboratory, Hampton Road, Teddington, Middlesex TW11 0LW, UK
[e] VSL, Thijsseweg 11, 2629 JA Delft, The Netherlands

ARTICLE INFO

Keywords:
Errors-in-variables
Straight-line regression
Weighted total least-squares
Law of propagation of uncertainty
Monte Carlo method
Implicit measurement model

ABSTRACT

Following the Guide to the expression of uncertainty in measurement (GUM), the slope and intercept in straight-line regression tasks can be estimated and their uncertainty evaluated by defining a measurement model. Minimizing the weighted total least-squares functional appropriately defines such a model when both regression input quantities ($\boldsymbol{X}$ and $\boldsymbol{Y}$) are uncertain.

This paper compares the uncertainty of the straight line evaluated by propagating distributions and by the law of propagation of uncertainty (LPU). The latter is in turn often approximated because the non-linear measurement model does not have closed form. We reason that the uncertainty recommended in the dedicated technical specification ISO/TS 28037:2010 does not fully implement the LPU (as intended) and can understate the uncertainty. A systematic simulation study quantifies this understatement and the circumstances where it becomes relevant. In contrast, the LPU uncertainty may often be appropriate. As a result, it is planned to revise ISO/TS 28037:2010.

1. Introduction

Estimating the straight-line relationship between two quantities is a fundamental task in metrology and other disciplines. For example, ensuring metrological traceability may involve the calibration of observations against a measurement standard over a certain range. Usually both quantities, say $\boldsymbol{X}$ and $\boldsymbol{Y}$, are uncertain and often the relationship between them is described by a straight line or even close to the identity line. Without evaluating the uncertainty, tasks such as calibrations, validations and performance evaluations are incomplete in metrology.

This research focuses on straight-line errors-in-variables regression, that is on estimating a straight line when the two quantities determining the regression line are both uncertain. More precisely, the focus will be on evaluating the uncertainty of the slope and intercept in straight-line errors-in-variables regression.

The Guide to the expression of uncertainty in measurement (in short GUM, [1–3]) builds on the formulation of a measurement model to derive estimates, associated uncertainties and possibly distributions for the output quantities. The uncertainties of the input quantities of this model and possible correlations between them must be known in advance. The measurement model relates the input quantities ($\boldsymbol{X}$ and $\boldsymbol{Y}$) to the output quantities (slope and intercept here), and for regression problems its definition depends on the estimation method [4–6]. In straight-line errors-in-variables regression, various methods are available to estimate the slope and intercept, such as least-squares, finite-sample adjusted, inequality-constrained or structured least-squares, least median of squares, maximum likelihood, method of moments and Bayesian methods (e.g. [7–13]). Even more algorithms exist to determine their solution. Different measurement models could thus be formulated for these regression problems. In Section 2.1, we will define the measurement model based on the weighted total least-squares (WTLS) method, which is the appropriate least-squares method (e.g. [14–17]) and is recommended in ISO publications [16,18]. Evaluating uncertainties based on the measurement model is regarded as the GUM perspective here and commonplace for regression applications in metrology. Alternative statistical approaches deriving asymptotic, resampling, subjective or other uncertainties will not be the focus.

Following the GUM, the uncertainty of measurands can be evaluated by applying the (linear) law of propagation of uncertainty (LPU) to the input quantities, i.e. by employing [1] or generalizing it to multiple and possibly implicitly defined output quantities [3]. This process of propagating uncertainties involves evaluating the partial derivatives of the measurement model at the estimates of the input quantities. It

* Corresponding author.
E-mail address: katy.klauenberg@ptb.de (K. Klauenberg).

https://doi.org/10.1016/j.measurement.2021.110340
Received 22 July 2021; Received in revised form 16 September 2021; Accepted 12 October 2021
Available online 6 November 2021

is described in Section 2.2 for WTLS-based measurement models, and explicit, closed form and approximate approaches are reviewed briefly. Alternatively, the uncertainty of measurands and other summary information can be evaluated by propagating probability distributions of input quantities, mainly by applying the Monte Carlo (MC) method in [2] or generalizing it to multiple output quantities [3], as described in Section 2.3.

Since the WTLS-based measurement model is generally not linear in the input quantities, it is not evident how different the uncertainties according to the LPU are compared to those according to MC. Moreover, the evaluation of uncertainties for the slope and intercept according to the LPU is difficult for the general WTLS-based measurement model, as derivatives of an implicit model and matrix inversions are involved [19]. Approximating the uncertainties according to the LPU is prevalent (cf. Section 2.2), and this research will scrutinize the uncertainty recommended in the technical specification ISO/TS 28037:2010 [16], which is based on the Jacobian of the residuals. Which properties does this uncertainty have? Is it close to the uncertainty evaluated according to the LPU? And is the latter in turn close to the uncertainty evaluated by applying MC? Section 3 will answer the above questions for the most common WTLS cases.

The ISO technical specification [16] addresses straight-line calibration problems having very general classes of uncertainty structure and offers algorithms for their solution. It was the first time such a comprehensive capability had been offered in an ISO publication. The technical specification claims to evaluate uncertainties according to the LPU even for the most general WTLS-based measurement model (see clauses 5.5.1 and 8.2.2 as well as Note 7 in clauses 7.2.1 and 10.2.2 in [16]). Section 3.1 pinpoints the theoretical relationship between the uncertainty evaluated according to the LPU and according to [16]. It shows that these uncertainties differ from each other — for the general model just as when explicit or closed formulae exist. While for the most general classes of calibration problems, the technical specification [16] acknowledges in clause 5.5.1 that its methods will be more accurate for data having small associated uncertainties, this paper shows that they are not a faithful implementation of the LPU. To ensure that the LPU is fully implemented, the (second-derivative) Hessian matrix must be used in addition to the Jacobian (first-derivative matrix) of the residuals and, moreover, the latter matrix should be evaluated at the observed rather than the fitted (modelled) points.

Section 3.2 then quantifies the difference between the uncertainties according to the LPU, the approximated LPU in [16] and MC for common WTLS-based models. An extensive simulation study for settings where the uncertainty according to the LPU can be evaluated explicitly shows that it is roughly $170\,K\,\%$ larger for the slope than the uncertainty evaluated according to the technical specification [16]. The latter is roughly $230\,K\,\%$ smaller than the uncertainty derived by MC, where K is the minimum of the variance in either input quantity divided by the variance of its fitted values. The simulation shows that the ISO/TS 28037:2010 routinely understates the uncertainty in straight-line errors-in-variables regression by several percent.

In Section 4 we draw conclusions for metrology and briefly discuss alternatives to measurement-model-based inference.

2. Evaluating uncertainty for the weighted total least-squares method

2.1. The measurement model

Let $\boldsymbol{X} = (X_1, \ldots, X_N)^\top$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_N)^\top$ be the two quantities determining the straight line and $\boldsymbol{x} = (x_1, \ldots, x_N)^\top$, $\boldsymbol{y} = (y_1, \ldots, y_N)^\top$ the estimates available for them. In addition, let $\boldsymbol{U}$ be the symmetric, positive definite covariance matrix containing the known uncertainties and correlations associated with $\boldsymbol{x}$ and $\boldsymbol{y}$. That is, the diagonal $\mathrm{diag}(\boldsymbol{U}) = \left(u^2(x_1), \ldots, u^2(x_N), u^2(y_1), \ldots, u^2(y_N)\right)^\top$ contains the variances. The off-diagonal elements $\boldsymbol{U}_{i,j} = r_{i,j}u(x_i)u(x_j)$ as well as $\boldsymbol{U}_{N+i,N+j} = r_{N+i,N+j}u(y_i)u(y_j)$ contain the correlations $r_{i,j}$ and $r_{N+i,N+j}$ among the quantities, and the elements $\boldsymbol{U}_{i,N+j} = r_{i,N+j}u(x_i)u(y_j)$ comprise the correlations $r_{i,N+j}$ between them for $i \neq j$ and $i, j = 1, \ldots, N$. Repeated observations are not considered.

Straight-line errors-in-variables regression then aims at estimating the intercept β_0 and slope β_1 best fitting the observations $\boldsymbol{x}, \boldsymbol{y}$ while considering all uncertainties and covariances in $\boldsymbol{U}$. The appropriate least-squares method is based on the WTLS functional

$$Q = \begin{pmatrix} \boldsymbol{x} - \boldsymbol{x}^* \\ \boldsymbol{y} - \boldsymbol{y}^* \end{pmatrix}^\top \boldsymbol{U}^{-1} \begin{pmatrix} \boldsymbol{x} - \boldsymbol{x}^* \\ \boldsymbol{y} - \boldsymbol{y}^* \end{pmatrix} \tag{1}$$

(e.g. [14–17]), which describes a sum of weighted squared residuals between the observations and the unknown true values $\boldsymbol{x}^*, \boldsymbol{y}^*$. The latter are linearly related:

$$\boldsymbol{y}^* = \beta_0 \mathbf{1} + \beta_1 \boldsymbol{x}^*,$$

where $\mathbf{1}$ denotes a vector of ones. The vector minimizing functional (1) with respect to β_0, β_1 and $\boldsymbol{x}^*$ defines the solution $(\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\boldsymbol{x}^*}^\top)$ of the WTLS method.

For straight-line errors-in-variables regression, the measurement model required by the GUM [1–3] can then be defined based on the WTLS method by replacing the estimates $\boldsymbol{x}$ and $\boldsymbol{y}$ in the minimization of Q by the underlying quantities $\boldsymbol{X} = (X_1, \ldots, X_N)^\top$ and $\boldsymbol{Y} = (Y_1, \ldots, Y_N)^\top$, respectively. (C.f. e.g. [5,15,17,20].) That is,

$$\left(\beta_0, \beta_1, \boldsymbol{X}^{*\top}\right)^\top = \underset{\widetilde{\beta}_0, \widetilde{\beta}_1, \widetilde{\boldsymbol{x}}}{\arg\min}\; Q(\boldsymbol{X}, \boldsymbol{Y}), \quad \text{with} \tag{2}$$
$$Q(\boldsymbol{X}, \boldsymbol{Y}) = \begin{pmatrix} \boldsymbol{X} - \widetilde{\boldsymbol{x}} \\ \boldsymbol{Y} - \widetilde{\beta}_0 \mathbf{1} - \widetilde{\beta}_1 \widetilde{\boldsymbol{x}} \end{pmatrix}^\top \boldsymbol{U}^{-1} \begin{pmatrix} \boldsymbol{X} - \widetilde{\boldsymbol{x}} \\ \boldsymbol{Y} - \widetilde{\beta}_0 \mathbf{1} - \widetilde{\beta}_1 \widetilde{\boldsymbol{x}} \end{pmatrix}.$$

Usually the intercept β_0 and the slope β_1 are of primary interest. Then (β_0, β_1) define the measurand and $\boldsymbol{X}^*$ are auxiliary output quantities.

We will call Eq. (2) the WTLS-based (measurement) model here. The model is multivariate, and usually non-linear and implicit. That is, model (2) is linear in $\boldsymbol{X}$ and $\boldsymbol{Y}$ only when, for example, the uncertainty in one of the two input quantities is zero, where it reduces to a weighted least-squares (WLS)-based measurement model. In addition, the model is not available as a closed form expression in $\boldsymbol{X}$ and $\boldsymbol{Y}$ except for special cases. Table 1 lists some well-known special WTLS cases and properties of their measurement models. Many other cases [21] and names are in use. For example, WTLS regression is also called generalized least-squares regression [15], generalized Gauss–Markov regression [21], or confusingly, just total least squares (TLS). Also case C may not be called Deming regression in some communities.

The general WTLS-based model requires iterative procedures to derive the estimates $\widehat{\beta}_0, \widehat{\beta}_1$ (see e.g. [14] for a review, [24] for numerical aspects or the technical specification [16]). The estimated straight line is invariant when interchanging the quantities $\boldsymbol{X}$ and $\boldsymbol{Y}$ (due to the symmetry of model (2) except for $\beta_1 = 0$) and $\widehat{\beta}_0, \widehat{\beta}_1$ are maximum likelihood estimators under normality assumptions [14].

Let us now focus on evaluating uncertainties associated with the slope and intercept of the straight line under measurement model (2). Following the GUM [3], this evaluation can be carried out by a linear propagation of uncertainties (described below) or by a propagation of distributions (Section 2.3).

2.2. Uncertainty according to the law of propagation of uncertainty

The LPU implies that uncertainties of all input quantities are propagated through a linearized measurement model [1,3]. Since the WTLS-based model (2) is non-linear in general, the uncertainties derived according to the LPU only approximate the uncertainty.

Let $\boldsymbol{U}^{\mathrm{LPU}}(\boldsymbol{X}^*, \beta_0, \beta_1)$ be the covariance matrix of the output derived by the LPU, i.e. containing the squared standard uncertainties and covariances after linearizing model (2). This matrix is derived via the

Table 1
Overview of some well-known special WTLS cases and the properties of their measurement models.

Case	Name	Definition	Measurement model
A	WLS	$u(x_i) = 0$ for all i	linear (Appendix A)
B	orthogonal regression, also known as TLS	$\boldsymbol{U} = c\boldsymbol{I}$	explicit
C	Deming regression	$r_{ij} = 0,\ u(y_i) = cu(x_i)$ for $i \neq j$	explicit (Eqs. (9) and (10))
D	cross correlation only	$r_{ij} = 0$ for all $i \neq j + N$	implicit, 1-d optimization [22,23]

implicit function theorem because the measurement model (2), also restated as

$$\frac{\partial Q(\boldsymbol{X},\boldsymbol{Y})}{\partial(\widetilde{\boldsymbol{x}}^\top,\widetilde{\beta}_0,\widetilde{\beta}_1)} = \mathbf{0},$$

generally depends on the input $\boldsymbol{X},\boldsymbol{Y}$ and the output $\beta_0,\beta_1,\boldsymbol{X}^*$. Formally, the covariance matrix can be expressed as

$$\boldsymbol{U}^{\mathrm{LPU}}(\boldsymbol{X}^*,\beta_0,\beta_1) = \left(\partial_{\beta\beta}Q\right)^{-1}\left(\partial_{\beta X}Q\right)\boldsymbol{U}\left(\partial_{\beta X}Q\right)^\top\left(\partial_{\beta\beta}Q\right)^{-1} \tag{3}$$

(see clause 6.3 in [3] for implicit models, [25] for minimization-based and [18,26] for errors-in-variables regression), where

$$\partial_{\beta\beta}Q := \frac{\partial^2 Q(\boldsymbol{X},\boldsymbol{Y})}{\partial(\widetilde{\boldsymbol{x}}^\top,\widetilde{\beta}_0,\widetilde{\beta}_1)\,\partial(\widetilde{\boldsymbol{x}}^\top,\widetilde{\beta}_0,\widetilde{\beta}_1)^\top} \quad \text{and}$$

$$\partial_{\beta X}Q := \frac{\partial^2 Q(\boldsymbol{X},\boldsymbol{Y})}{\partial(\widetilde{\boldsymbol{x}}^\top,\widetilde{\beta}_0,\widetilde{\beta}_1)\,\partial(\boldsymbol{X}^\top,\boldsymbol{Y}^\top)^\top}$$

denote the Hessian of $Q(\boldsymbol{X},\boldsymbol{Y})$ with respect to the output and the Jacobian of the model with respect to the input quantities, respectively. All derivatives are evaluated at the estimates $\widehat{\boldsymbol{x}^*},\widehat{\beta}_0,\widehat{\beta}_1$ and $\boldsymbol{x},\boldsymbol{y}$. An algorithm to evaluate (3) exactly up to numerical errors is described in [19,27] for WTLS-based models without cross correlation, and this could be extended to the general case.

The covariance matrix (3) can be given in closed form for special cases only. For cases B to D in Table 1 (or clauses 7 and 8 in the technical specification [16]), [22,28] state the measurement model for the slope as an implicit 'cubic' function, apply the implicit function theorem and derive the LPU variances as

$$u^2(\beta_1)^{\mathrm{LPU}} = \frac{\sum_i W_i^2\left(U_i^2u^2(y_i) + V_i^2u^2(x_i) - 2r_{i,n+i}V_iU_iu(x_i)u(y_i)\right)}{D^2}, \tag{4}$$

$$u^2(\beta_0)^{\mathrm{LPU}} = \frac{1}{\sum_i W_i} + \left(2\bar{\widehat{x}}^* - \bar{x}\right)^2u^2(\beta_1)^{\mathrm{LPU}} + \frac{2\left(2\bar{\widehat{x}}^* - \bar{x}\right)\left(\bar{\widehat{x}}^* - \bar{x}\right)}{D}, \quad \text{with} \tag{5}$$

$$D = 1/\widehat{\beta}_1\sum_i W_iU_iV_i + 4\sum_i W_iU_i^*\left(\widehat{x}_i^* - x_i\right) - 1/\widehat{\beta}_1\sum_i W_i^2(\widehat{\beta}_1U_i - V_i)^2r_{i,n+i}u(x_i)u(y_i),$$

where $W_i = \left(\widehat{\beta}_1^2u^2(x_i) + u^2(y_i) - 2\widehat{\beta}_1r_{i,n+i}u(x_i)u(y_i)\right)^{-1}$, $\bar{z} = \left(\sum_i W_iz_i\right)/\left(\sum_i W_i\right)$ denotes the weighted average of a vector, $V_i = y_i - \bar{y}$, $U_i = x_i - \bar{x}$ and $U_i^* = \widehat{x}_i^* - \bar{\widehat{x}}^*$. For constant variance ratios without correlation (cases B and C in Table 1), $\widehat{\beta}_1$ can be stated explicitly (see Section 3.2), the uncertainties $u(\beta_1)^{\mathrm{LPU}}, u(\beta_0)^{\mathrm{LPU}}$ (termed LPU uncertainties here) become explicit as well, and the covariance is easily calculated:

$$u(\beta_0,\beta_1)^{\mathrm{LPU}} = -\bar{x}u^2(\beta_1)^{\mathrm{LPU}}.$$

For TLS-based models (case B in Table 1 assuming equal uncertainties and no correlation) whose line passes through the origin, [25] derived the exact LPU uncertainty for the angle of the line (see sec. 7 and equ. (38) in [25]), which can be transformed to an exact LPU uncertainty of the slope (see Appendix C).

For the WTLS-based model with cross correlation only (case D in Table 1) and an angle parametrization of the line, [23,29] derive the inverse Hessian. However, the implicit dependence of the solution on the input quantities is not accounted for (see Appendix D).

The literature provides additional approximations of the LPU uncertainty, and [19,30] review some of them. For general WTLS-based measurement models, the technical specification ISO/TS 28037:2010 [16] and also [15] recommend the uncertainty

$$\boldsymbol{U}^{\mathrm{ISO}}(\boldsymbol{X}^*,\beta_0,\beta_1) = \left(\boldsymbol{J}^\top\boldsymbol{U}^{-1}\boldsymbol{J}\right)^{-1}, \quad \text{with } \boldsymbol{J} = \begin{pmatrix} -\boldsymbol{I} & 0 & 0 \\ -\widetilde{\beta}_1\boldsymbol{I} & -\mathbf{1} & -\widetilde{\boldsymbol{x}}^* \end{pmatrix} \tag{6}$$

being the Jacobian of the unweighted residuals. However, how does this prevalent approximation compare to the application of the LPU for the general WTLS-based model? Is the difference important? The next section compares the LPU uncertainty with the approximation in the technical specification. We do so theoretically in Section 3.1 and quantify the differences empirically in Section 3.2.

Recall that the LPU uncertainty in turn is only an approximation of the true uncertainty. The linearization may lead to smaller or larger uncertainties. Section 3.2 will empirically quantify this difference as well.

2.3. Uncertainty according to the Monte Carlo method

The GUM provides an alternative to approximating uncertainties by the LPU, namely the propagation of distributions, usually via the MC method [3]. While the LPU requires estimates and uncertainties, but no distributional assumptions, the MC method inherently does. In return, the MC method provides a joint distribution for the output, i.e. more information than just estimates, standard uncertainties and correlation.

We follow the GUM (clause 6.4.8 in [2]) and assign a multivariate Normal distribution

$$\begin{pmatrix}\boldsymbol{X}\\ \boldsymbol{Y}\end{pmatrix} \sim N\left(\begin{pmatrix}\boldsymbol{x}\\ \boldsymbol{y}\end{pmatrix},\boldsymbol{U}\right) \tag{7}$$

to the input quantities, assuming that no other knowledge is available except for the input estimates, their standard uncertainties and correlations. Different assumptions lead to different distributions (such as in [31]) and to conclusions to be researched in future but are suspected to differ from ours. The joint distribution of the input quantities is then propagated through the WTLS-based measurement model (2) applying MC. That is, new observations $(\boldsymbol{x}'^\top,\boldsymbol{y}'^\top)^\top$ are sampled from (7) repeatedly and model (2) is applied to each. The resulting solutions $\left(\beta_0',\beta_1'\right)$ and possibly $\boldsymbol{x}^{*\prime}$ are then samples from the joint distribution of output quantities and form a discrete representation thereof (clause 7.5 in [3]). Depending on the number of samples drawn for the input quantities, the samples of the output quantities approximate the distribution arbitrarily closely. Consequently, the mean and (co)variance of the output sample (c.f. clause 7.6 in [3]) can be used to approximate the estimate and its uncertainty. Likewise, additional summary information may be derived from the sample, such as coverage intervals or regions. Under the assumption that any differences due to drawing only a finite sample are small, we will treat the uncertainties derived by MC as the reference here. MC methods are seen, at least by Bayesian statisticians, as a 'gold standard' for the numerical calculation of uncertainties [32].

For WTLS-based measurement models, the MC method is also described in [17,31,33]. In contrast, [19,22,34] draw samples centred around the output estimates $\widehat{\boldsymbol{x}^*},\widehat{\boldsymbol{y}^*}$ instead of the input estimates $\boldsymbol{x},\boldsymbol{y}$, which describes a kind of bootstrap [35]. Even though [36] draws samples around $\boldsymbol{x},\boldsymbol{y}$, a WLS-based measurement model is used. These simulations do not implement the MC method described in the GUM [2,3] for the WTLS-based model (2) and hence are not used for comparisons here.

3. Comparing uncertainty approximations for the weighted total least-squares method

3.1. Theoretical properties of linearized uncertainties

Let us compare the exact LPU uncertainty (3) with its prevalent approximation in the technical specification ISO/TS 28037:2010 [16]. The latter recommends $\boldsymbol{U}^{\mathrm{ISO}}(\boldsymbol{X}^*, \beta_0, \beta_1)$ in (6) as the covariance matrix for the slope and intercept in errors-in-variables regressions, based on the Jacobian $\boldsymbol{J}$ of the unweighted residuals $\boldsymbol{r}$, evaluated at the estimates $(\widehat{\boldsymbol{x}^*}^\top, \widehat{\beta}_0, \widehat{\beta}_1)$. We can observe a list of properties of $\boldsymbol{U}^{\mathrm{ISO}}(\boldsymbol{X}^*, \beta_0, \beta_1)$ compared to $\boldsymbol{U}^{\mathrm{LPU}}(\boldsymbol{X}^*, \beta_0, \beta_1)$.

Firstly, when one of the quantities $\boldsymbol{X}, \boldsymbol{Y}$ is known exactly (case A in Table 1), the uncertainty derived in the technical specification coincides with the LPU uncertainty, as one would expect. (See Appendix A.)

Secondly, the approximated uncertainty is equivalent to evaluating the exact LPU uncertainty at the fitted points $\widehat{\boldsymbol{x}^*}, \widehat{\boldsymbol{y}^*}$ instead of at the observed ones $\boldsymbol{x}, \boldsymbol{y}$. In [22] this was shown for case D in Table 1, where the uncertainty in Eqs. (4) and (5) reduces to

$$u^2(\beta_1)^{\mathrm{ISO}} = \frac{1}{\sum_i W_i \left(\hat{x}_i^* - \bar{\hat{x}}^*\right)^2}, \quad u^2(\beta_0)^{\mathrm{ISO}} = \frac{1}{\sum_i W_i} + \left(\bar{\hat{x}}^*\right)^2 u^2(\beta_1)^{\mathrm{ISO}} \tag{8}$$

and is identical to the uncertainties recommended in clause 8 of the technical specification [16]. Appendix B extends this finding to the general WTLS-based model and shows that replacing $\boldsymbol{x}, \boldsymbol{y}$ in the covariance matrix (3) with $\widehat{\boldsymbol{x}^*}, \widehat{\boldsymbol{y}^*}$ reduces to the matrix (6).

Furthermore, the covariance matrix (6) can be viewed as an additional linearization of the LPU covariance matrix (3) in β_1 and $\boldsymbol{X}^*$. That is, if the unweighted residuals $\boldsymbol{r}$ were linear in $\beta_1, \boldsymbol{X}^*$, then $\partial_{\beta\beta} Q$ would collapse into $\boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{J}$ (see (B.1) in Appendix B) and thus $\boldsymbol{U}^{\mathrm{ISO}}(\boldsymbol{X}^*, \beta_0, \beta_1) = \boldsymbol{U}^{\mathrm{LPU}}(\boldsymbol{X}^*, \beta_0, \beta_1)$. However, $\boldsymbol{r}$ is not linear in β_1 or $\boldsymbol{X}^*$.

Along the same lines, the uncertainty recommended in technical specification [16] can be viewed as the uncertainty of the step $\Delta = \left(\boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{J}\right)^{-1} \boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{r}$ performed by the Gauss–Newton minimization algorithm (c.f. equ. (4) in [25]). This counter-intuitive dependence of the uncertainty on the minimization algorithm was observed in [25] and can be solved by replacing the step direction $\left(\boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{J}\right)^{-1}$ by the inverse Hessian $2\left(\partial_{\beta\beta} Q\right)^{-1}$.

In a nutshell, the uncertainty evaluated in the technical specification [16] differs from the LPU uncertainty. This difference will be large whenever $\boldsymbol{x}$ or $\boldsymbol{y}$ are far from $\widehat{\boldsymbol{x}^*}$ or $\widehat{\boldsymbol{y}^*}$, or when $\beta_1 \boldsymbol{X}^*$ is markedly non-linear. Since the data $\boldsymbol{x}, \boldsymbol{y}$ are more variable than the fitted values $\widehat{\boldsymbol{x}^*}, \widehat{\boldsymbol{y}^*}$, we suspect [16] to underrate the LPU uncertainty. For TLS-based models (case B in Table 1) with independent and identical normally distributed residuals, this underrating was shown in [25, sec. 6].

Empirically, [19,27] observed differences for the uncertainties of the slope and intercept for a single data set and [37], for 10 data sets. We will expand these observations in a systematic simulation study.

3.2. Empirical properties – Simulations

Simulations will empirically quantify the differences between the uncertainties according to MC, the LPU and the approximated LPU as in [16]. We examine case C in Table 1, assuming no correlation and constant variance ratios $u^2(y_i) = c u^2(x_i)$. This setting has the advantage that explicit formulas for the estimates of the slope and intercept as well as their uncertainties are available. That is, the comparison of uncertainties will not be masked by numerical errors due to an iterative algorithm for the estimates. In particular, the estimates are calculated following [38], with the slope and intercept being

$$\widehat{\beta}_1 = \frac{B}{2\sum_i U_i V_i u^{-2}(x_i)} + \frac{\left(B^2 + 4c\left(\sum_i U_i V_i u^{-2}(x_i)\right)^2\right)^{1/2}}{2\sum_i U_i V_i u^{-2}(x_i)} \tag{9}$$

$$\widehat{\beta}_0 = \bar{y} - \widehat{\beta}_1 \bar{x}, \quad \text{where } B = \sum_i V_i^2 u^{-2}(x_i) - c \sum_i U_i^2 u^{-2}(x_i), \tag{10}$$

and their uncertainties are evaluated according to the LPU (Eqs. (4), (5)) as well as the approximated LPU (Eq. (8)). In addition, the estimates and uncertainties were evaluated from 10^5 MC draws according to (7), where for each MC draw, the estimates were calculated. See Appendix E for validations of the programming code. Systematic simulation studies for more complex settings are desirable, but are computationally more demanding and probably need to be restricted to (application) specific settings due to the sheer variety.

The simulation study covers small as well as large data sets and regressions with a variety of different slopes. The uncertainty in x covers a range from tiny to considerable, and either the uncertainty of both quantities is roughly equal to $u(x) \approx u(y)$ or $u(x)|\beta_1| \approx u(y)$, or either uncertainty dominates. In particular, the simulation includes all combinations of $\beta_1, u^2(y) = u^2(y_i)$ and N listed in Table 2 and a grid of values $u^2(x) = u^2(x_i)$ smaller than $0.25\mathrm{Var}(\boldsymbol{x}^*)$. For each setting in Table 2, N true values x_i^* were randomly drawn from $U(0, 1)$. Without loss of generality, the intercept $\beta_0 = 0$ did not vary during simulations (but was estimated). One data set $\boldsymbol{x}, \boldsymbol{y}$ was then drawn from the Normal distribution

$$\begin{pmatrix} \boldsymbol{x} \\ \boldsymbol{y} \end{pmatrix} \sim N\left(\begin{pmatrix} \boldsymbol{x}^* \\ \beta_0 \mathbf{1} + \beta_1 \boldsymbol{x}^* \end{pmatrix}, \boldsymbol{U}\right). \tag{11}$$

For comparison purposes, 100 data sets were simulated for the setting $N = 500$, $\beta_1 = 1$, $u(y) = 0.1$, $u(x)/\mathrm{sd}(\boldsymbol{x}^*) = 0.2$ and a 95% interval of the results has a half-width below 0.0071 in Fig. 1 and 0.0047 in Fig. 2, where sd denotes the standard deviation and Var the variance).

The difference between the estimates derived from the exact formula (9) and from the MC draws is tiny compared with the difference to the simulated values and compared with the uncertainty. (Cf. Figure F.1 in the supplement.) Estimates are not considered further.

Fig. 1 displays on the left-hand side the differences between the standard uncertainty for the slope evaluated according to the LPU (formula (4)) and the approximated LPU as in [16] (formula (8)). On the right-hand side Fig. 1 displays the differences between the uncertainty evaluated according to MC (see Section 2.3) and the approximated LPU as in [16] (formula (8)). The ratios between these uncertainties are displayed for data sets of size $N = 500$ and discussed subsequently. The supplement displays these ratios for fewer observations of $N = 100$ and $N = 20$.

Let us first quantify the approximated LPU uncertainty in terms of the LPU uncertainty. The LPU uncertainty is always larger than its approximation, and their ratio is roughly given by

$$\frac{u(\beta_1)^{\mathrm{LPU}}}{u(\beta_1)^{\mathrm{ISO}}} \approx 1 + 1.7\min(K), \quad \text{with } K = \left(\frac{u^2(x)}{\mathrm{Var}\left(\hat{\boldsymbol{x}}^*\right)}, \frac{u^2(y)}{\mathrm{Var}\left(\hat{\boldsymbol{y}}^*\right)}\right) \tag{12}$$

(displayed by the pink and dashed lines in the left graph of Fig. 1). This empirical approximation works well when roughly $u(y) \notin (1/2, 2)\, u(x)|\beta_1|$ and when N is large. For small N, data sets from the same simulation setting will differ more, and we observe more variability around the approximate ratio (12); compare Figure F.3 in the supplement for $N = 20$. For $u(y) \approx u(x)\beta_1$ (marked by the open circles in Fig. 1), the approximated LPU uncertainty will be slightly closer to the exact LPU uncertainty than in (12). Altogether, the uncertainty in the technical specification [16] will be more than 5% smaller than the LPU uncertainty it seeks to approximate, whenever $\min\left(u(x), u(y)/|\beta_1|\right) / \sqrt{\mathrm{Var}\left(\hat{\boldsymbol{x}}^*\right)} > 0.2$ (dotted lines in Fig. 1). This difference is due to the technical specification [16] evaluating the involved derivatives at the fitted data points instead of the observed ones (see Section 3.1).

Moreover, we quantify the approximated LPU uncertainty in [16] in terms of the uncertainty evaluated by MC. The MC uncertainty is larger than the approximated LPU uncertainty, and their ratio is roughly given by

$$\frac{u(\beta_1)^{\mathrm{MC}}}{u(\beta_1)^{\mathrm{ISO}}} \approx 1 + 2.3\min(K), \quad \text{with } K = \left(\frac{u^2(x)}{\mathrm{Var}\left(\hat{\boldsymbol{x}}^*\right)}, \frac{u^2(y)}{\mathrm{Var}\left(\hat{\boldsymbol{y}}^*\right)}\right) \tag{13}$$

Table 2
Overview of all combinations of the slope β_1, the uncertainty $u^2(y)$ and the number of observations N for the simulation study. For each such setting and for a grid of uncertainties $u^2(x) \leq 0.021$, one data set $(\mathbf{x}^\top, \mathbf{y}^\top)$ was drawn at random from (11).

$u^2(y)$	$u(y)/\mathrm{sd}(\mathbf{x}^*)$	$\beta_1 = -1$	$\beta_1 = 0.1$	$\beta_1 = 1$	$\beta_1 = 2$	$\beta_1 = 10$
0.01	0.35	$N = 100$	20, 100, 500	100, 500	100, 500	20, 100, 500
0.05	0.77	$N = 100$	20, 100, 500	100, 500	100	20, 100
0.1	1.1	$N = 100$	20, 100, 500	20, 100, 500	20, 100, 500	20, 100, 500
0.2	1.5	$N = 100$	20, 100, 500	100, 500	100	20, 100

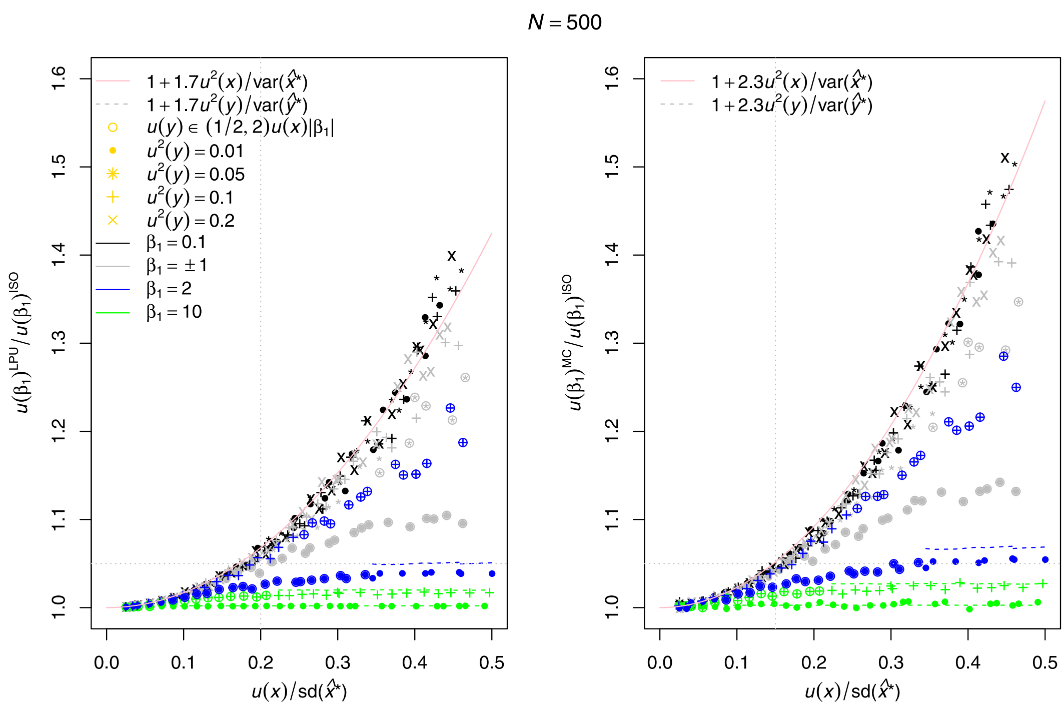

Fig. 1. Ratios of the standard uncertainty for the slope evaluated according to the LPU and the approximated LPU as in [16] (left), as well as evaluated according to MC and the approximated LPU (right) for $N = 500$ observations and for the settings in Table 2. See text for details.

(displayed by the pink and dashed lines in the right graph of Fig. 1). Again, this empirical approximation works well when $u(y) \notin (1/2, 2)\,u(x)|\beta_1|$ and when N is large. Small numbers of observations (such as $N \approx 20$) need further examination, as the ratio $u(\beta_1)^{\mathrm{MC}}/u(\beta_1)^{\mathrm{ISO}}$ seems to be better approximated by $1 + 2.3\min(K) + 0.2\max(K)$ when $\min(K) = u^2(y)/\mathrm{Var}\left(\hat{\mathbf{y}}^*\right)$, cf. Figure F.3 in the supplement. In addition, for small N and noisy data sets, the number of MC draws may need to exceed 10^5 by far.[1] Altogether, the uncertainty in the technical specification [16] will be more than 5% smaller than the MC uncertainty, whenever $\min\left(u(x), u(y)/|\beta_1|\right) / \sqrt{\mathrm{Var}\left(\hat{\mathbf{x}}^*\right)} > 0.15$ (dotted lines in Fig. 1). This difference is due to the technical specification [16] linearizing the measurement model and evaluating the involved derivatives at the fitted data points instead of the observed ones (see Section 3.1).

Likewise, the ratio of the uncertainty evaluated by MC and by the LPU is displayed in Fig. 2 (left) and can be approximated. The LPU uncertainty is usually smaller and will be more than 5% smaller than the MC uncertainty only when $\min\left(u(x), u(y)/|\beta_1|\right) / \sqrt{\mathrm{Var}\left(\hat{\mathbf{x}}^*\right)} > 0.4$. To ease comparison, all three ratios $u(\beta_1)^{\mathrm{MC}}/u(\beta_1)^{\mathrm{ISO}}$, $u(\beta_1)^{\mathrm{LPU}}/u(\beta_1)^{\mathrm{ISO}}$ and $u(\beta_1)^{\mathrm{MC}}/u(\beta_1)^{\mathrm{LPU}}$ are displayed in Fig. 2 (right) for the setting $\beta_1 = 1, u^2(y) = 0.1$ and $N = 500$. While propagating uncertainties through a linearized model may approximate the uncertainty of the slope well, the approximated LPU as in [16] shows marked discrepancies for most of these data sets.

These conclusions and the results shown in Figs. 1, 2, F.2, F.3 are representative for arbitrary values of the intercept β_0 and for values of $N, \beta_1, u(y)$ in the range of those in Table 2 as well as for values $u(x)$ up to half the standard deviation of the true values $\boldsymbol{x}^*$. The randomness of each simulated data set contributes little. In addition, the results scale to settings with other ranges of true values, as first simulations for $x^* \sim U(0,2)$ and $x^* \sim N(1, 1/12)$ indicate. Furthermore, simulating data sets with $\beta_1 = 1, N = 500$ and varying squared uncertainties $u^2(y_i) \sim N(0.1, 0.02^2)$ for a range of variance ratios $c = u^2(y_i)/u^2(x_i) \leq 0.21$ shows very similar results compared to the data sets with $u^2(y) = 0.1$. It is therefore likely that the above results extend to the general setting C in Table 1 with explicit formulas for estimates and uncertainties.

4. Discussion and conclusions

This research reviewed how straight-line errors-in-variables regression can be approached by closely following the internationally recognized GUM. When the uncertainties of the two quantities involved in the regression are known, it is straightforward and standard to base the measurement model on the weighted total least-squares functional. Estimates and uncertainties for the slope and intercept can then be evaluated either by propagating estimates and uncertainties of the input quantities through the linearized model or by propagating their distributions through the full model.

[1] We observed unreasonably large MC uncertainty ratios, i.e. $u(\hat{\beta}_1)^{\mathrm{MC}} > 3u(\hat{\beta}_1)^{\mathrm{ISO}}$, for $N = 20$, $\beta = 0.1$, few $u^2(x)/\mathrm{Var}\left(\hat{\boldsymbol{x}}^*\right) \approx 0.45^2$ and all simulated values $u(y)$. These 6 out of 360 data sets are not displayed in Figure F.3 in the supplement.

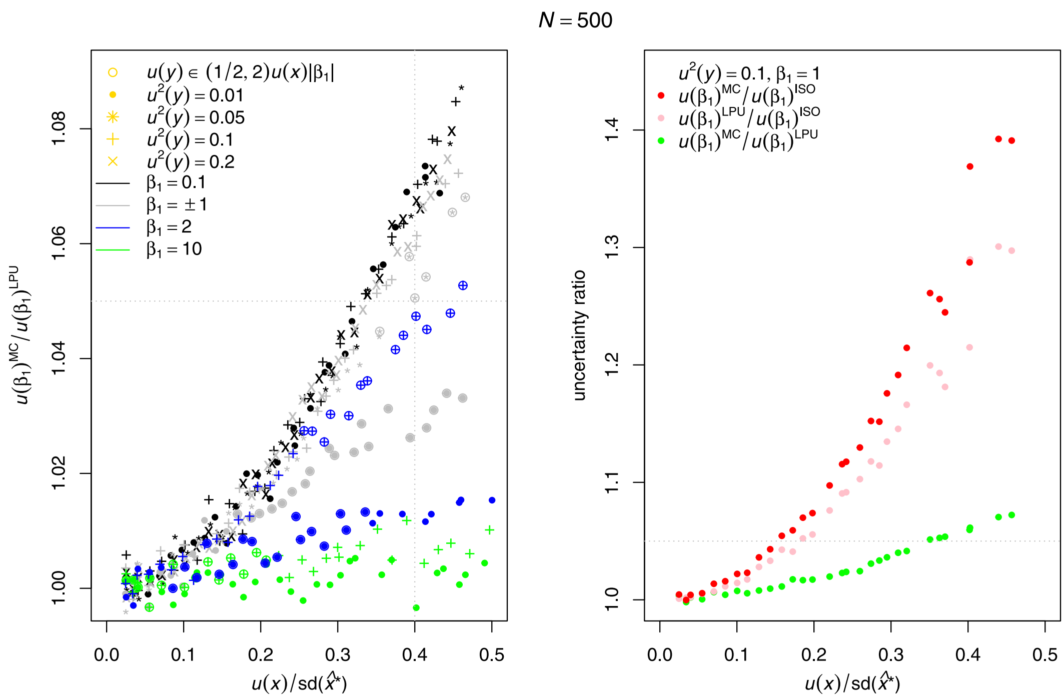

Fig. 2. Ratios of the standard uncertainty for the slope evaluated according to the LPU and MC for $N = 500$ observations and for the settings in Table 2 (left graph). For comparison, all three ratios for the uncertainty of the slope (evaluated according to the LPU, approximated LPU and MC) are displayed to the right for a single simulation setting ($\beta_1 = 1, u^2(y) = 0.1, N = 500$).

While the latter requires a numerical approach, namely the Monte Carlo method, the analytic expressions for the former are generally difficult. They include the evaluation and derivatives of an implicit function as well as matrix inversion methods. The uncertainties evaluated by the (linear) law of propagation of uncertainty differ from those evaluated by propagating distributions because the weighted total least-squares based measurement model is non-linear. This research is the first to quantify this difference in a systematic simulation study. In particular, for uncorrelated input quantities each with constant variance, the uncertainty for the slope evaluated by the linear law will be more than 5% smaller than the Monte Carlo uncertainty when both input uncertainties are larger than 0.4 times the spread of their fitted values. Such settings may be rare in metrology, but could occur, for instance, when comparing two methods with similar, considerable uncertainty. The difference between propagating distributions and uncertainties may also become relevant when an uncertainty underrated by 3% to 4% is unacceptable.

In addition, the uncertainties are often not evaluated directly by the linear law but approximated again. This research focused on the uncertainty based on the Jacobian of the residuals, as recommended in the technical specification ISO/TS 28037:2010 dedicated to straight-line fitting. For the general weighted total least-squares based measurement model, it was shown that the covariance matrix from the linear law is only equal to its approximation if the involved derivatives were evaluated at the fitted data points instead of the observed ones. The uncertainties recommended in the technical specification therefore differ from ones evaluated by the law of propagation of uncertainty, which was shown for special cases before. For the first time, the difference was quantified in a systematic simulation study. In particular, for uncorrelated input quantities each with constant variance, we empirically determined the ratio of the approximation to the uncertainties from the linear law and to the Monte Carlo uncertainties for the slope. The uncertainty recommended in the technical specification will underrate the true uncertainty by more than 5% when both input uncertainties are larger than 0.15 times the spread of their fitted values. For many of these settings, uncertainties from the linear law will still be appropriate.

Weighted total least squares is devised particularly to estimate the straight-line relationship when the uncertainty for both quantities is non-negligible; and for these settings, the discrepancy of the uncertainty following ISO/TS 28037:2010 from the true uncertainty is pronounced.

Propagating uncertainties by the linear law seems to be adequate for the weighted total least-squares based measurement model, at least for most metrological straight-line settings without correlation and constant variance ratios. Correlation, other uncertainty structures and linear settings occur in practice and should be accounted for (e.g. [39–41]). The current simulation study can only caution against evaluating the uncertainty following ISO/TS 28037:2010, unless the range of either input quantity is large compared to its uncertainty. We recommend revising the technical specification or limiting its scope. In general, we recommend evaluating uncertainties of the straight line for weighted total least-squares based models by propagating distributions.

The technical specification ISO/TS 28037:2010 evaluates the derivative of the measurement model at the fitted points. Future research may explore the general relationship of such an approach to statistical inferences. That is, instead of basing uncertainty evaluations on a measurement model, a statistical model with distributional assumptions would be formulated, which is described amongst other places in the recent, less restrictive GUM document [33] (especially clause 11.4). For errors-in-variables models, distributions are then assigned to the residuals and estimators can be derived according to various principles. Maximum likelihood is one of these principles and the Bayesian method, a further one. The Bayesian approach is flexible and attractive when additional information is available prior to collecting the data. For errors-in-variables models, [6,17,42] describe such an approach for the metrology community. However, the choice of a prior distribution for the fitted values seems to be problematic when no information is available [13,43,44].

CRediT authorship contribution statement

Katy Klauenberg: Conceptualization, Formal analysis, Funding acquisition, Investigation, Methodology, Software, Validation, Visualization, Writing – original draft, Writing – review & editing. **Steffen Martens:** Conceptualization, Formal analysis, Investigation, Software, Validation, Writing – review & editing. **Alen Bošnjaković:** Validation, Writing – review & editing. **Maurice G. Cox:** Funding acquisition, Project administration, Conceptualization, Validation, Writing – review & editing. **Adriaan M.H. van der Veen:** Validation, Writing – review & editing. **Clemens Elster:** Funding acquisition, Conceptualization, Supervision, Validation, Visualization, Writing – review & editing.

Declaration of competing interest

One or more of the authors of this paper have disclosed potential or pertinent conflicts of interest, which may include receipt of payment, either direct or indirect, institutional support, or association with an entity in the biomedical field which may be perceived to have potential conflict of interest with this work. For full disclosure statements refer to https://doi.org/10.1016/j.measurement.2021.110340. This work is part of the project 17NRM05 EMUE, which has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme.

Acknowledgments

The authors would like to thank Alistair Forbes (NPL) for his feedback on the manuscript and valuable discussions on wider interpretations of the GUM. This work is part of the project 17NRM05 EMUE, which has received funding from the EMPIR programme co-financed by the Participating States and from the European Union's Horizon 2020 research and innovation programme.

Appendix A. The weighted least-squares (WLS) case

When the uncertainty in one of the input quantities, say $\boldsymbol{X}$, vanishes and any cross correlation with it, the WLS estimate

$$(\hat{\beta}_0, \hat{\beta}_1)^\top = \left(\boldsymbol{D}^\top \boldsymbol{U}_y^{-1} \boldsymbol{D}\right)^{-1} \boldsymbol{D}^\top \boldsymbol{U}_y^{-1} \boldsymbol{y}$$

coincides with the WTLS estimate, where $\boldsymbol{D} = (\boldsymbol{1}, \boldsymbol{x}^\top)$ denotes the Jacobian or design matrix and $\boldsymbol{U}_y$, the covariance matrix for $\boldsymbol{y}$. The derived measurement model is a linear function in its input $\boldsymbol{Y}$, and thus the covariance matrix of the slope and intercept is known to be

$$\begin{aligned}\boldsymbol{U}(\beta_0, \beta_1) &= \left(\boldsymbol{D}^\top \boldsymbol{U}_y^{-1} \boldsymbol{D}\right)^{-1} \boldsymbol{D}^\top \boldsymbol{U}_y^{-1} \boldsymbol{U}_y \boldsymbol{U}_y^{-1} \boldsymbol{D} \left(\boldsymbol{D}^\top \boldsymbol{U}_y^{-1} \boldsymbol{D}\right)^{-1} \\ &= \left(\boldsymbol{D}^\top \boldsymbol{U}_y^{-1} \boldsymbol{D}\right)^{-1}.\end{aligned}$$

Not surprisingly, this covariance matrix coincides with the implicit formula (3) as $\partial_{\beta X} Q = -2\boldsymbol{D}^\top \boldsymbol{U}_y^{-1}$ and $\partial_{\beta\beta} Q = 2\boldsymbol{D}^\top \boldsymbol{U}_y^{-1} \boldsymbol{D}$, which can also be inferred from [25, sec. 5] with $Q(\boldsymbol{Y}) = \epsilon^\top \boldsymbol{U}_y^{-1} \epsilon$ and $\epsilon = \boldsymbol{Y} - \boldsymbol{D}\,(\widetilde{\beta}_0, \widetilde{\beta}_1)^\top$.

Appendix B. The Jacobian-based uncertainty

The algorithm in clause 10 of the technical specification ISO/TS 28037:2010 [16] approximates the covariance matrix (3) according to the LPU by

$$\boldsymbol{U}^{\mathrm{ISO}}(\boldsymbol{X}^*, \beta_0, \beta_1) = \left(\boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{J}\right)^{-1},$$

where $\boldsymbol{J} = \frac{\partial \boldsymbol{r}}{\partial(\widetilde{\boldsymbol{x}}, \widetilde{\beta}_0, \widetilde{\beta}_1)} = \begin{pmatrix} -\boldsymbol{I} & 0 & 0 \\ -\widetilde{\beta}_1 \boldsymbol{I} & \boldsymbol{-1} & -\widetilde{\boldsymbol{x}}^* \end{pmatrix}$ is the Jacobian of the unweighted residuals $\boldsymbol{r}$ that is evaluated at the estimates $\widehat{\boldsymbol{x}^*}, \hat{\beta}_0, \hat{\beta}_1$, and where $\boldsymbol{r}^\top = \left(\boldsymbol{X}^\top - \widetilde{\boldsymbol{x}}^\top,\ \boldsymbol{Y}^\top - \left(\widetilde{\beta}_0 \boldsymbol{1} + \widetilde{\beta}_1 \widetilde{\boldsymbol{x}}\right)^\top\right)$ such that $Q(\boldsymbol{X}, \boldsymbol{Y}) = \boldsymbol{r}^\top \boldsymbol{U}^{-1} \boldsymbol{r}$.

The parts of the covariance matrix $\boldsymbol{U}^{\mathrm{LPU}}(\boldsymbol{X}^*, \beta_0, \beta_1)$ in (3) can be calculated as

$$\begin{aligned}\partial_{\beta\beta} Q &= 2\frac{\partial \boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{r}}{\partial(\widetilde{\boldsymbol{x}}, \widetilde{\beta}_0, \widetilde{\beta}_1)^\top} = 2\frac{\partial \boldsymbol{J}^\top}{\partial(\widetilde{\boldsymbol{x}}, \widetilde{\beta}_0, \widetilde{\beta}_1)^\top} \boldsymbol{U}^{-1} \boldsymbol{r} + 2\boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{J} \\ \partial_{\beta X} Q &= 2\frac{\partial \boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{r}}{\partial(\boldsymbol{X}, \boldsymbol{Y})^\top} = 2\boldsymbol{J}^\top \boldsymbol{U}^{-1}\end{aligned} \tag{B.1}$$

because $\boldsymbol{J}$ is constant in the data but not in $\widetilde{\boldsymbol{x}}^*$ and $\widetilde{\beta}_1$. If we now evaluate the covariance matrix (3) at $\widehat{\boldsymbol{x}^*}, \widehat{\boldsymbol{y}^*}$ and $\widehat{\boldsymbol{x}^*}, \hat{\beta}_0, \hat{\beta}_1$, instead of at the estimates $\boldsymbol{x}, \boldsymbol{y}$ and $\widehat{\boldsymbol{x}^*}, \hat{\beta}_0, \hat{\beta}_1$, the residuals $\boldsymbol{r}$, and with them the first part of $\partial_{\beta\beta} Q$, vanish. The covariance matrix (3) would then collapse to the matrix $\left(\boldsymbol{J}^\top \boldsymbol{U}^{-1} \boldsymbol{J}\right)^{-1}$ recommended in clause 10 of the technical specification [16].

That is, the technical specification [16] can be viewed as evaluating the LPU uncertainty at the fitted instead of at the observed values, which extends the observations in [22] to the general WTLS-based model.

Appendix C. Linearization of a reparametrized implicit function

Let a be the one-dimensional quantity whose linearly approximated standard uncertainty is $u(a)$ and let $Q(a)$ be the implicit functional whose solution $\partial_a Q = 0$ gives the estimate $\hat{a}$. In addition, let $a = f(\alpha)$ be some differentiable reparametrization and $Q(\alpha) = Q(f(\alpha))$, the reparametrized functional. We will show that the standard uncertainty $u(a)$ derived from linearizing $Q(a)$ is equivalent to first deriving the standard uncertainty $u(\alpha)$ from linearizing $Q(\alpha)$ and then linearly transforming α to a.

Let x be one of the input quantities of $\partial_a Q$. The contribution of x to the linearly approximated squared standard uncertainty (3) is given by

$$u_x^2(a) = \left(\partial_{ax} Q\right)^2 u^2(x) \left(\partial_{aa} Q\right)^{-2},$$

using the abbreviation for derivatives as in Section 2.2.

Alternatively, the standard uncertainty can be calculated from the linearly approximated standard uncertainty $u(\alpha)$ and linearizing f. Let this uncertainty be called $v(a)$. Then the contribution of x to $v^2(a)$ is

$$\begin{aligned}v_x^2(a) &= u_x^2(\alpha) \left(\partial_\alpha f\right)^2 \\ &= \left(\partial_{\alpha x} Q\right)^2 u^2(x) \left(\partial_{\alpha\alpha} Q\right)^{-2} \left(\partial_\alpha f\right)^2 \\ &= \left(\partial_{ax} Q \partial_\alpha f + \cancel{\partial_a Q \partial_{\alpha x} f}\right)^2 u^2(x) \left(\partial_{aa} Q \left(\partial_\alpha f\right)^2 + \cancel{\partial_a Q \partial_{\alpha\alpha} f}\right)^{-2} \left(\partial_\alpha f\right)^2 \\ &= u_x^2(a).\end{aligned}$$

Similarly, the standard uncertainty with respect to further input quantities can be derived: $v_y(a) = u_y(a)$.

Assuming no correlation between input quantities, the linearization of f and $Q(\alpha)$ in the input quantities is thus equivalent to directly linearizing $Q(a)$.

Appendix D. The Hessian-based uncertainty

The measurement model in [23,29] is reduced to an implicit formula for the angle parameter α, which depends neither on the second parameter nor on the fitted values (cf. (44) and (46) in [23]). For this functional Q, the mixed derivative with respect to α and the second parameter will be zero, and thus the uncertainty of the angle will be $u^2(\alpha)^{\mathrm{H}} = 2\left(\partial_{\alpha\alpha} Q\right)^{-1}$ according to the authors.

According to the LPU however, the squared standard uncertainty is

$$\begin{aligned}u^2(\alpha)^{\mathrm{LPU}} = &\sum_i \left[\left(\frac{\partial^2 Q}{\partial\alpha\, \partial X_i}\right)^2 u^2(x_i) + \left(\frac{\partial^2 Q}{\partial\alpha\, \partial Y_i}\right)^2 u^2(y_i)\right] \left(\partial_{\alpha\alpha} Q\right)^{-2} \\ &+ 2\sum_i \left[\frac{\partial^2 Q}{\partial\alpha\, \partial X_i} \frac{\partial^2 Q}{\partial\alpha\, \partial Y_i} u(x_i, y_i)\right] \left(\partial_{\alpha\alpha} Q\right)^{-2},\end{aligned}$$

derived from the one-dimensional version of (3) with block-diagonal structure $\boldsymbol{U}$, which will not be identical to $u^2(\alpha)^{\mathrm{H}}$ for arbitrary input uncertainties.

Appendix E. Validation of the programming code

The estimate (9) and the LPU uncertainty (4) were validated against the results from the CCC software release 1.3 [19] for one data set with the setting $N = 20, \beta_1 = 1, u^2(y) = 0.1, u(x)/\mathrm{sd}(\boldsymbol{x}^*) = 0.21$. The digits displayed by the software all agree with our results. For this data set, the LPU uncertainty is 14% larger than the approximated LPU uncertainty. The same data set was used to validate the approximated LPU uncertainty (8) against an independent implementation following the algorithm in clause 10 of the technical specification [16]. The first four significant digits agree.

Appendix F. Supplementary data

Supplementary material related to this article can be found online at https://doi.org/10.1016/j.measurement.2021.110340.

References

[1] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, Evaluation of measurement data – Guide to the expression of uncertainty in measurement, in: Joint Committee for Guides in Metrology, JCGM 100:2008, International Organization for Standardization, Geneva, 2008.

[2] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, Evaluation of measurement data – Supplement 1 to the Guide to the expression of uncertainty in measurement – Propagation of distributions using a Monte Carlo method, in: Joint Committee for Guides in Metrology, JCGM 101: 2008, International Organization for Standardization, Geneva, 2008.

[3] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, Evaluation of measurement data – Supplement 2 to the Guide to the expression of uncertainty in measurement – Extension to any number of output quantities, in: Joint Committee for Guides in Metrology, JCGM 102: 2011, International Organization for Standardization, Geneva, 2011.

[4] K. Klauenberg, G. Wübbeler, B. Mickan, P. Harris, C. Elster, A tutorial on Bayesian normal linear regression, Metrologia (ISSN: 0026-1394) 52 (6) (2015) 878, http://dx.doi.org/10.1088/0026-1394/52/6/878.

[5] C. Elster, Bayesian uncertainty analysis compared with the application of the GUM and its supplements, Metrologia (ISSN: 0026-1394) 51 (4) (2014) S159, http://dx.doi.org/10.1088/0026-1394/51/4/S159.

[6] A. Possolo, B. Toman, Assessment of measurement uncertainty via observation equations, Metrologia 44 (6) (2007) 464–475, http://dx.doi.org/10.1088/0026-1394/44/6/005.

[7] W.A. Fuller, in: N. Hoboken (Ed.), Measurement Error Models, first ed., Wiley-Interscience, ISBN: 978-0-470-09571-3, 2006.

[8] P. Xu, J. Liu, C. Shi, Total least squares adjustment in partial errors-in-variables models: algorithm and statistical analysis, J. Geod. 86 (8) (2012) 661–675.

[9] X. Fang, On non-combinatorial weighted total least squares with inequality constraints, J. Geod. 88 (8) (2014) 805–816.

[10] X. Fang, A structured and constrained total least-squares solution with cross-covariances, Studia Geophy. Et Geod. 58 (1) (2014) 1–16.

[11] X. Fang, W. Zeng, Y. Zhou, B. Wang, On the total least median of squares adjustment for the pattern recognition in point clouds, Measurement 160 (2020) 107794, URL https://doi.org/10.1016/j.measurement.2020.107794.

[12] P. Gustafson, Measurement Error and Misclassification in Statistics and Epidemiology: Impacts and Bayesian Adjustments, Chapman and Hall/CRC, ISBN: 978-1-58488-335-7, 2003.

[13] A. Zellner, An Introduction to Bayesian Inference in Econometrics, Wiley-Interscience, New York, ISBN: 978-0-471-16937-6, 2010.

[14] I. Markovsky, S. Van Huffel, Overview of total least-squares methods, Signal Process. 87 (10) (2007) 2283–2302.

[15] M.J.T. Milton, P.M. Harris, I.M. Smith, A.S. Brown, B.A. Goody, Implementation of a generalized least-squares method for determining calibration curves from data with general uncertainty structures, Metrologia (ISSN: 0026-1394) 43 (4) (2006) S291, URL https://doi.org/10.1088/0026-1394/43/4/S17.

[16] ISO/TS, 28037:2010 Determination and Use of Straight-Line Calibration Functions, International Organization for Standardization (ISO), Geneva, 2010.

[17] I. Lira, D. Grientschnig, Error-in-variables models in calibration, Metrologia (ISSN: 0026-1394) 54 (6) (2017) S133, http://dx.doi.org/10.1088/1681-7575/aa8f02.

[18] ISO, 6143:2001 Gas Analysis—Comparison Methods for Determining and Checking the Composition of Calibration Gas Mixtures, International Organization for Standardization (ISO), Geneva, 2001.

[19] A. Malengo, F. Pennecchi, A weighted total least-squares algorithm for any fitting model with correlated variables, Metrologia (ISSN: 0026-1394) 50 (6) (2013) 654, http://dx.doi.org/10.1088/0026-1394/50/6/654.

[20] C. Elster, B. Toman, Bayesian uncertainty analysis for a regression model versus application of GUM supplement 1 to the least-squares estimate, Metrologia (ISSN: 0026-1394) 48 (5) (2011) 233, http://dx.doi.org/10.1088/0026-1394/48/5/001.

[21] M.G. Cox, A.B. Forbes, P.M. Harris, I.M. Smith, The Classification and Solution of Regression Problems for Calibration, NPL Report CMSC 24/03, 2014.

[22] D. York, N.M. Evensen, M.L. Martínez, J. De Basabe Delgado, Unified equations for the slope, intercept, and standard errors of the best straight line, Amer. J. Phys. (ISSN: 0002-9505) 72 (3) (2004) 367–375, http://dx.doi.org/10.1119/1.1632486.

[23] M. Krystek, M. Anton, A least-squares algorithm for fitting data points with mutually correlated coordinates to a straight line, Meas. Sci. Technol. (ISSN: 0957-0233) 22 (3) (2011) 035101, http://dx.doi.org/10.1088/0957-0233/22/3/035101.

[24] G.H. Golub, C.F.V. Loan, Matrix Computations, in: Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, ISBN: 978-1-4214-0794-4, 2013.

[25] A. Balsamo, G. Mana, F. Pennecchi, The expression of uncertainty in non-linear parameter estimation, Metrologia 43 (5) (2006) 396–402, http://dx.doi.org/10.1088/0026-1394/43/5/009.

[26] W. Bremser, W. Hässelbarth, Controlling uncertainty in calibration, Anal. Chim. Acta 348 (1–3) (1997) 61–69.

[27] M. Lecuna, F. Pennecchi, A. Malengo, P.G. Spazzini, Calibration curve computing (CCC) software v2.0: A new release of the INRIM regression tool, Meas. Sci. Technol. 31 (11) (2020) 114004, http://dx.doi.org/10.1088/1361-6501/ab7d6e.

[28] D. York, Least squares fitting of a straight line with correlated errors, Earth Planet Sci. Lett. (ISSN: 0012-821X) 5 (1969) 320–324, http://dx.doi.org/10.1016/S0012-821X(68)80059-7.

[29] M. Krystek, M. Anton, A weighted total least-squares algorithm for fitting a straight line, Meas. Sci. Technol. 18 (11) (2007) 3438, URL https://doi.org/10.1088/0957-0233/18/11/025.

[30] C.A. Cantrell, Technical note: Review of methods for linear least-squares fitting of data and application to atmospheric chemistry problems, Atmospheric Chem. Phys. (ISSN: 1680-7316) 8 (17) (2008) 5477–5487, http://dx.doi.org/10.5194/acp-8-5477-2008.

[31] F.R. Guenther, A. Possolo, Calibration and uncertainty assessment for certified reference gas mixtures, Anal. Bioanal. Chem. (399) (2011) 489–500.

[32] J. Huggins, M. Kasprzak, T. Campbell, T. Broderick, Validated variational inference via practical posterior error bounds, in: S. Chiappa, R. Calandra (Eds.), Proceedings of the Twenty Third International Conference on Artificial Intelligence and Statistics, in: Proceedings of Machine Learning Research, 108, PMLR, 2020, pp. 1792–1802, URL http://proceedings.mlr.press/v108/huggins20a.html.

[33] BIPM, IEC, IFCC, ILAC, ISO, IUPAC, IUPAP, OIML, Guide to the expression of uncertainty in measurement – Part 6: Developing and using measurement models, in: Joint Committee for Guides in Metrology, JCGM GUM-6: 2020, International Organization for Standardization, Geneva, 2020.

[34] T. Bartel, S. Stoudt, A. Possolo, Force calibration using errors-in-variables regression and Monte Carlo uncertainty evaluation, Metrologia (ISSN: 0026-1394) 53 (3) (2016) 965–980, http://dx.doi.org/10.1088/0026-1394/53/3/965.

[35] R.J. Tibshirani, B. Efron, An Introduction to the Bootstrap, in: Monographs on Statistics and Applied Probability, Vol. 57, Chapman and Hall, New York, 1993.

[36] I. Lira, Monte Carlo evaluation of the uncertainty associated with the construction and use of a fitted curve, Measurement 44 (10) (2011) 2156–2164, URL https://doi.org/10.1016/j.measurement.2011.07.021.

[37] D.M. Titterington, A.N. Halliday, On the fitting of parallel isochrons and the method of maximum likelihood, Chem. Geol. 26 (3–4) (1979) 183–195, URL https://doi.org/10.1016/0009-2541(79)90045-7.

[38] D. York, Least-squares fitting of a straight line, Can. J. Phys. (ISSN: 0008-4204) 44 (5) (1966) 1079–1086, http://dx.doi.org/10.1139/p66-090.

[39] S. Martens, K. Klauenberg, B. Mickan, C. Yardin, N. Fischer, C. Elster, EMUE-D4-3-QuantifyUncertaintiesInCalibration, Zenodo, 2020, http://dx.doi.org/10.5281/ZENODO.4016915.

[40] S. Martens, K. Klauenberg, J. Neukammer, S. Cowen, S.L.R. Ellison, C. Elster, EMUE-D5-4-MethodComparisonWithCorrelation, Zenodo, 2020, http://dx.doi.org/10.5281/ZENODO.3911583.

[41] K. Snow, Topics in Total Least-Squares Adjustment Within the Errors-In-Variables Model: Singular Cofactor Matrices and Prior Information, (Ph.D. thesis), The Ohio State University, 2012.

[42] A.M.H. van der Veen, Bayesian methods for type A evaluation of standard uncertainty, Metrologia (ISSN: 0026-1394) 55 (5) (2018) 670, http://dx.doi.org/10.1088/1681-7575/aad103.

[43] J.-P. Florens, M. Mouchart, J.-F. Richard, BayesIan inference in error-in-variables models, J. Multivariate Anal. (ISSN: 0047-259X) 4 (4) (1974) 419–452, http://dx.doi.org/10.1016/0047-259X(74)90023-2.

[44] D. Leonard, Estimating a bivariate linear relationship, Bayesian Anal. 6 (4) (2011) 727–754, http://dx.doi.org/10.1214/11-BA627.