AUSTRIAN JOURNAL OF STATISTICS
Volume 35 (2006), Number 2&3, 397–406

Scheffé-type Confidence Region for the Calibration Line

Gejza Wimmer[1] and Viktor Witkovský[2]
[1]Masaryk University, Brno, Czech Republic
[2] Slovak Academy of Sciences, Bratislava, Slovak Republic

Abstract: Analyzed is the comparative calibration problem in the case when linear relationship is assumed between two considered measuring devices. The method deriving the approximate confidence region for the unknown parameters of the calibration line based on estimation of the calibration line via replicated errors-in-variables model is given in Wimmer et al. (2004). Essential point in this approach is the use of the F-approximation of the distribution of the F-statistic suggested by Kenward and Roger (1997). Moreover, in Wimmer and Witkovský (2005) such a calibration procedure is conducted which enables proper interval estimates of several unknown values based on the next independent measurement results obtained by the less precise measuring device.

The present contribution is a continuation of work presented in Wimmer et al. (2004) and Wimmer and Witkovský (2005). Derived is the Scheffé-type confidence region for the (whole) calibration line. This enables to construct the interval estimators for the multiple-use calibration case.

Keywords: Comparative Univariate Linear Calibration, Kenward-Roger Type Approximation.

1 The Calibration Problem

Under the term *calibration problem* we understand the task of fitting the calibration curve based on well designed calibration experiment. The calibration curve expresses the relationship between the ideal (true, errorless) results of measuring the same object (substance, quantity) by two measuring devices. In particular, we are interested in finding the proper estimators of the coefficients of the calibration curve and constructing the interval estimators for further determinations of the unknown true value of the measured substance in units of the more precise instrument, given measurement in units of a less precise instrument.

In this paper we consider comparative calibration, the situation in which one instrument is calibrated against the other and both are subject to the measurement error. We suggest a method for constructing the Scheffé-type confidence region for the whole calibration line. This could be directly used for the multiple-use calibration. For more details on classical approaches to the calibration problem see e.g. Eisenhart (1939), Krutchkoff (1967), Scheffé (1973), and Brown (1993).

Throughout this paper we will consider the above mentioned problem under the following circumstances:

(i) The less precise measuring device, say instrument A: The measurement result x_i obtained by the instrument A is a realization of normally distributed random variable

X_i, i.e. $X_i \sim N(\mu_i, \sigma_x^2)$, $i = 1, \ldots, n$, where the mean value $\mathrm{E}(X_i) = \mu_i$ is the ideal (true, errorless) value of the measure and given in units of the instrument A and σ_x^2 is the (unknown) dispersion of X_i assumed to be common for all i.

(ii) The more precise measuring device, say instrument B: The measurement result y_i obtained by the the instrument B is a realization of normally distributed random variable Y_i, i.e. $Y_i \sim N(\nu_i, \sigma_y^2)$, $i = 1, \ldots, n$, where the mean value $\mathrm{E}(Y_i) = \nu_i$ is the ideal (true, errorless) value of the measure and given in units of the instrument B and σ_y^2 is the (unknown) dispersion of Y_i assumed to be common for all i.

(iii) For estimation of the parameters of the calibration curve and for obtaining the confidence region of the parameters we accomplish a pre-planned calibration experiment with measurements made by both instruments, A and B, respectively, on a set of $n \geq 4$ suitably chosen objects (substances, quantities of interest), say $v_1, \ldots, v_n$, repeatedly $m \geq 2$ times for each object.

(iv) All the measurements are mutually independent.

(v) We assume that over the typical range of values of μ and ν (the range of interest) the true, however unknown, the calibration curve is a linear function, i.e. $\nu = a + b\mu$ with (unknown) parameters a, b.

In the above setup the instrument B is considered to be more precise than the instrument A ($\sigma_y^2 \leq \sigma_x^2$). The instrument A is said to be the calibrated device.

2 Estimation of the Calibration Line

Wimmer et al. (2004) suggested an iterative algorithm for estimating the parameters of the linear calibration curve of the following comparative calibration model, see also Kubáček and Kubáčková (2000).

Denote the vector of errorless measurement results made by the instrument A by $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)'$ and the vector of errorless measurement results made by the instrument B by $\boldsymbol{\nu} = (\nu_1, \ldots, \nu_n)'$, respectively. Let the measurements made by the instrument A be represented by n-dimensional random vector $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \sigma_x^2 \boldsymbol{I}_{n,n})$ and let the measurements made by the instrument B be represented by $\boldsymbol{Y} \sim N(\boldsymbol{\nu}, \sigma_y^2 \boldsymbol{I}_{n,n})$, where $\boldsymbol{I}$ represents the identity matrix.

Assuming (i)-(v) (first, ignoring the replications of measurements) we have the following calibration model

$$\begin{pmatrix} \boldsymbol{X} \\ \boldsymbol{Y} \end{pmatrix} \sim N\left[\begin{pmatrix} \boldsymbol{\mu} \\ \boldsymbol{\nu} \end{pmatrix}, \begin{pmatrix} \sigma_x^2 \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_y^2 \boldsymbol{I} \end{pmatrix} \right] \tag{1}$$

with nonlinear constraints on parameters

$$\boldsymbol{\nu} = a\boldsymbol{1} + b\boldsymbol{\mu}\,, \tag{2}$$

where $\boldsymbol{1} = \boldsymbol{1}_{n,1} = (1, \ldots, 1)'$ is an n-dimensional vector of ones, and a and b are unknown coefficients which specify the intercept and the slope of the calibration line.

Further, we suggest to linearize the model (1) and (2) by using Taylor series expansion locally about $\boldsymbol{\mu}_0 = (\mu_{01}, \ldots, \mu_{0n})'$ and b_0 (some values chosen near to the true parameters

$\boldsymbol{\mu}$ and b). So, $\boldsymbol{\mu} = \boldsymbol{\mu}_0 + \delta\boldsymbol{\mu}$, $b = b_0 + \delta b$ and the new parameters of the approximate linear model are $\delta\boldsymbol{\mu} = (\delta\mu_1, \ldots, \delta\mu_n)'$, $\boldsymbol{\nu}$, a, δb, σ_x^2, σ_y^2:

$$\begin{pmatrix} \boldsymbol{X} - \boldsymbol{\mu}_0 \\ \boldsymbol{Y} \end{pmatrix} \sim N\left[\begin{pmatrix} \delta\boldsymbol{\mu} \\ \boldsymbol{\nu} \end{pmatrix}, \begin{pmatrix} \sigma_x^2 \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \sigma_y^2 \boldsymbol{I} \end{pmatrix} \right] \tag{3}$$

with linear constraints

$$b_0\boldsymbol{\mu}_0 + (b_0\boldsymbol{I} \vdots -\boldsymbol{I}) \begin{pmatrix} \delta\boldsymbol{\mu} \\ \boldsymbol{\nu} \end{pmatrix} + (\boldsymbol{1}, \boldsymbol{\mu}_0) \begin{pmatrix} a \\ \delta b \end{pmatrix} = \boldsymbol{0}\,. \tag{4}$$

Model (3) – (4) is highly overparameterized, so we repeat the measurements on n objects independently m times. Let us denote by $\boldsymbol{\xi}_1, \ldots, \boldsymbol{\xi}_m$ the m independent replications of

$$\boldsymbol{\xi} = \begin{pmatrix} \boldsymbol{X} - \boldsymbol{\mu}_0 \\ \boldsymbol{Y} \end{pmatrix},$$

i.e. the vector $\underline{\boldsymbol{\xi}} = (\boldsymbol{\xi}_1', \ldots, \boldsymbol{\xi}_m')'$ represents the vector of all measurements from the calibration experiment. The linearized version of the replicated model (3) and (4) is then given by

$$\underline{\boldsymbol{\xi}} \sim N\left[(\boldsymbol{1}_{m,1} \otimes \boldsymbol{I}_{2n,2n}) \begin{pmatrix} \delta\boldsymbol{\mu} \\ \boldsymbol{\nu} \end{pmatrix}, \boldsymbol{I}_{m,m} \otimes (\sigma_x^2 \boldsymbol{V}_1 + \sigma_y^2 \boldsymbol{V}_2) \right], \tag{5}$$

with linear constraints (4) on the parameters, where

$$\boldsymbol{V}_1 = \begin{pmatrix} \boldsymbol{I}_{n,n} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{V}_2 = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{I}_{n,n} \end{pmatrix},$$

and by $\otimes$ we denote the Kronecker product of matrices.

Assuming that the model (5) with linear constraints (4) is (approximately) correct, the BLUE (best linear unbiased estimators) of the parameters $\boldsymbol{\mu} = \boldsymbol{\mu}_0 + \delta\boldsymbol{\mu}$, $\boldsymbol{\nu}$, a, and δb are derived in Wimmer et al. (2004) and are given by

$$\hat{\boldsymbol{\mu}} = \bar{\boldsymbol{X}} + \frac{b_0\sigma_x^2}{b_0^2\sigma_x^2 + \sigma_y^2} \boldsymbol{M}_{[1,\mu_0]} (\bar{\boldsymbol{Y}} - b_0\bar{\boldsymbol{X}})\,, \tag{6}$$

$$\hat{\boldsymbol{\nu}} = \bar{\boldsymbol{Y}} - \frac{\sigma_y^2}{b_0^2\sigma_x^2 + \sigma_y^2} \boldsymbol{M}_{[1,\mu_0]} (\bar{\boldsymbol{Y}} - b_0\bar{\boldsymbol{X}})\,, \tag{7}$$

$$\begin{pmatrix} \hat{a} \\ \hat{\delta b} \end{pmatrix} = \begin{pmatrix} n & \boldsymbol{1}'\boldsymbol{\mu}_0 \\ \boldsymbol{\mu}_0'\boldsymbol{1} & \boldsymbol{\mu}_0'\boldsymbol{\mu}_0 \end{pmatrix}^{-1} \begin{pmatrix} \boldsymbol{1}'(\bar{\boldsymbol{Y}} - b_0\bar{\boldsymbol{X}}) \\ \boldsymbol{\mu}_0'(\bar{\boldsymbol{Y}} - b_0\bar{\boldsymbol{X}}) \end{pmatrix} \tag{8}$$

with the covariance matrix

$$\operatorname{cov} \begin{pmatrix} \hat{a} \\ \hat{b} \end{pmatrix} = \frac{b_0^2\sigma_x^2 + \sigma_y^2}{m} \begin{pmatrix} n & \boldsymbol{1}'\boldsymbol{\mu}_0 \\ \boldsymbol{\mu}_0'\boldsymbol{1} & \boldsymbol{\mu}_0'\boldsymbol{\mu}_0 \end{pmatrix}^{-1},$$

where $\hat{b} = b_0 + \hat{\delta b}$ and $\bar{\boldsymbol{X}} = \frac{1}{m}\sum_{j=1}^m \boldsymbol{X}_j$, $\boldsymbol{X}_j = (X_{j1}, \ldots, X_{jn})'$, $\bar{\boldsymbol{Y}} = \frac{1}{m}\sum_{j=1}^m \boldsymbol{Y}_j$, $\boldsymbol{Y}_j = (Y_{j1}, \ldots, Y_{jn})'$, $j = 1, \ldots, m$, and $\boldsymbol{M}_{[1,\mu_0]} = \boldsymbol{I} - [\boldsymbol{1}, \boldsymbol{\mu}_0] \left([\boldsymbol{1}, \boldsymbol{\mu}_0]'[\boldsymbol{1}, \boldsymbol{\mu}_0]\right)^{-1} [\boldsymbol{1}, \boldsymbol{\mu}_0]'$.

The estimates of the parameters of the calibration line and covariance matrices depend on the unknown variance components σ_x^2 and σ_y^2, which could be suitably estimated from the data by the (iterated) $(\sigma_{x0}^2, \sigma_{y0}^2)$-MINQUE, i.e. the $(\sigma_{x0}^2, \sigma_{y0}^2)$-locally minimum norm quadratic unbiased estimator.

The $(\sigma_{x0}^2, \sigma_{y0}^2)$-MINQUE of $(\sigma_x^2, \sigma_y^2)'$ in linear model (5) with constraints (4) is given by

$$\begin{pmatrix} \hat\sigma_x^2 \\ \hat\sigma_y^2 \end{pmatrix} = \frac{1}{n(m-1)} \left[\boldsymbol{I}_{2,2} - c_0 \begin{pmatrix} b_0^4\sigma_{x0}^4 & b_0^2\sigma_{x0}^4 \\ b_0^2\sigma_{y0}^4 & \sigma_{y0}^4 \end{pmatrix} \right] \begin{pmatrix} \hat\kappa_1 \\ \hat\kappa_2 \end{pmatrix}, \tag{9}$$

where

$$\begin{aligned} c_0 &= \frac{n-2}{(b_0^4\sigma_{x0}^4 + \sigma_{y0}^4)(mn-2) + 2b_0^2\sigma_{x0}^2\sigma_{y0}^2(m-1)n}, \\ \hat\kappa_1 &= \sum_{j=1}^{m} (\boldsymbol{X}_j - \bar{\boldsymbol{X}})'(\boldsymbol{X}_j - \bar{\boldsymbol{X}}) + m(\bar{\boldsymbol{X}} - \hat{\boldsymbol{\mu}})'(\bar{\boldsymbol{X}} - \hat{\boldsymbol{\mu}}), \\ \hat\kappa_2 &= \sum_{j=1}^{m} (\boldsymbol{Y}_j - \bar{\boldsymbol{Y}})'(\boldsymbol{Y}_j - \bar{\boldsymbol{Y}}) + m(\bar{\boldsymbol{Y}} - \hat{\boldsymbol{\nu}})'(\bar{\boldsymbol{Y}} - \hat{\boldsymbol{\nu}}). \end{aligned}$$

The covariance matrix (correct locally at $(\sigma_{x0}^2, \sigma_{y0}^2)$) of the estimator (9) is

$$\boldsymbol{W} = \begin{pmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{pmatrix} = \frac{2}{n(m-1)} \left[\boldsymbol{I}_{2,2} - c_0 \begin{pmatrix} b_0^4\sigma_{x0}^4 & b_0^2\sigma_{x0}^4 \\ b_0^2\sigma_{y0}^4 & \sigma_{y0}^4 \end{pmatrix} \right] \begin{pmatrix} \sigma_{x0}^4 & 0 \\ 0 & \sigma_{y0}^4 \end{pmatrix}. \tag{10}$$

All the estimators strongly depend on the chosen initial values $\boldsymbol{\mu}_0$, b_0, σ_{x0}^2, σ_{y0}^2, and further, on the quality of linearization of the originally nonlinear model (1)–(2) at these initial values, which on the other hand strongly depends on the experimental design of the calibration experiment. If there is no specific prior information on the true values of the parameters, a natural choice of the initial values, estimated from the measured data, could be the following:

$$\begin{aligned} \hat{\boldsymbol{\mu}}_0 &= \bar{\boldsymbol{X}}, \qquad \hat b_0 = \frac{n\bar{\boldsymbol{X}}'\bar{\boldsymbol{Y}} - (\boldsymbol{1}'\bar{\boldsymbol{X}})(\boldsymbol{1}'\bar{\boldsymbol{Y}})}{n\bar{\boldsymbol{X}}'\bar{\boldsymbol{X}} - (\boldsymbol{1}'\bar{\boldsymbol{X}})^2}, \\ \hat\sigma_{x0}^2 &= \frac{1}{n(m-1)} \sum_{i=1}^{n}\sum_{j=1}^{m} (X_{ji} - \bar X_i)^2, \\ \hat\sigma_{y0}^2 &= \frac{1}{n(m-1)} \sum_{i=1}^{n}\sum_{j=1}^{m} (Y_{ji} - \bar Y_i)^2, \end{aligned} \tag{11}$$

with $\bar X_i = m^{-1}\sum_{j=1}^{m} X_{ij}$ and $\bar Y_i = m^{-1}\sum_{j=1}^{m} Y_{ij}$, $i = 1, \ldots, n$. Further we compute $\hat a$, $\hat b$ from (8), $\hat{\boldsymbol{\mu}}$ from (6), $\hat{\boldsymbol{\nu}}$ from (7), $\hat\sigma_x^2$ and $\hat\sigma_y^2$ from (9).

The estimation procedure could be iterated in such a way until convergence is reached according to the following algorithm:

1. Set the initial values $\hat{\boldsymbol{\mu}}^{(0)} = \hat{\boldsymbol{\mu}}_0$, $\hat b^{(0)} = \hat b_0$, $\hat\sigma_x^{2(0)} = \hat\sigma_{x0}^2$, and $\hat\sigma_y^{2(0)} = \hat\sigma_{y0}^2$, according to (11).
2. For given $\hat{\boldsymbol{\mu}}^{(k)}$ and $\hat b^{(k)}$ set $\boldsymbol{\mu}_0 = \hat{\boldsymbol{\mu}}^{(k)}$, $b_0 = \hat b^{(k)}$ and estimate $\hat a^{(k+1)}$, $\hat b^{(k+1)}$ from (8).

3. Using this $\hat{b}^{(k+1)}$ and $\hat{\boldsymbol{\mu}}^{(k)}$, $\hat{\sigma}_x^{2(k)}$, $\hat{\sigma}_y^{2(k)}$, set $b_0 = \hat{b}^{(k+1)}$, $\boldsymbol{\mu}_0 = \hat{\boldsymbol{\mu}}^{(k)}$, $\sigma_{x0}^2 = \hat{\sigma}_x^{2(k)}$, $\sigma_{y0}^2 = \hat{\sigma}_y^{2(k)}$ and estimate $\hat{\boldsymbol{\mu}}^{(k+1)}$ from (6) and $\hat{\boldsymbol{\nu}}^{(k+1)}$ from (7).
4. Finally, using $\hat{\sigma}_x^{2(k)}$, $\hat{\sigma}_y^{2(k)}$, and $\hat{b}^{(k+1)}$, $\hat{\boldsymbol{\mu}}^{(k+1)}$, $\hat{\boldsymbol{\nu}}^{(k+1)}$, set $\sigma_{x0}^2 = \hat{\sigma}_x^{2(k)}$, $\sigma_{y0}^2 = \hat{\sigma}_y^{2(k)}$, $b_0 = \hat{b}^{(k+1)}$, $\hat{\boldsymbol{\mu}} = \hat{\boldsymbol{\mu}}^{(k+1)}$, $\hat{\boldsymbol{\nu}} = \hat{\boldsymbol{\nu}}^{(k+1)}$, and estimate $\hat{\sigma}_x^{2(k+1)}$ and $\hat{\sigma}_y^{2(k+1)}$ from (9). Set $k = k + 1$ and go to the point 2.
5. After reaching convergence and finishing the iterations calculate the covariance matrix $\boldsymbol{W}$ according to (10).

For testing the null hypothesis $H_0 : (a, b)' = (a_*, b_*)'$ and for the construction of the confidence region for the parameters $(a, b)'$ we suggest to use the F-statistic

$$F = \frac{1}{2} \begin{pmatrix} \hat{a} - a_* \\ \hat{b} - b_* \end{pmatrix}' \hat{\Phi}^{-1} \begin{pmatrix} \hat{a} - a_* \\ \hat{b} - b_* \end{pmatrix}, \tag{12}$$

where

$$\hat{\Phi} = \frac{b_0^2 \hat{\sigma}_x^2 + \hat{\sigma}_y^2}{m} \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}_0 \\ \boldsymbol{\mu}_0'\mathbf{1} & \boldsymbol{\mu}_0'\boldsymbol{\mu}_0 \end{pmatrix}^{-1},$$

with the values of the parameters given from the last iteration of the algorithm.

Under H_0 the distribution of the F-statistic is approximated by the Fisher-Snedecor F-distribution with 2 and u degrees of freedom (for more details see Kenward and Roger, 1997), where

$$u = (mn - 2) + \frac{2n(m-1) b_0^2 \hat{\sigma}_x^2 \hat{\sigma}_y^2}{b_0^4 \hat{\sigma}_x^4 + \hat{\sigma}_y^4}, \tag{13}$$

with the values of the parameters estimated by the algorithm. For more details see Wimmer et al. (2004) and Wimmer and Witkovský (2005).

3 Scheffé-type Confidence Region for the Calibration Line

Let $(a, b)'$ represent the true vector of parameters of the calibration line $\nu = a + b\mu$, $\mu \in \langle \mu_l, \mu_u \rangle$, where the interval $\langle \mu_l, \mu_u \rangle$ represents the typical range of the calibration experiment.

Using the derivation in Section 2 the parameter estimates given by the algorithm, we have the following (approximate) distribution

$$F = \frac{1}{2} \frac{m}{b_0^2 \hat{\sigma}_x^2 + \hat{\sigma}_y^2} \begin{pmatrix} \hat{a} - a \\ \hat{b} - b \end{pmatrix}' \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}_0 \\ \boldsymbol{\mu}_0'\mathbf{1} & \boldsymbol{\mu}_0'\boldsymbol{\mu}_0 \end{pmatrix} \begin{pmatrix} \hat{a} - a \\ \hat{b} - b \end{pmatrix} \sim F_{2,u}, \tag{14}$$

with u given by (13). From that we get

$$\Pr\left\{ \frac{1}{2} \frac{m}{b_0^2 \hat{\sigma}_x^2 + \hat{\sigma}_y^2} \begin{pmatrix} \hat{a} - a \\ \hat{b} - b \end{pmatrix}' \begin{pmatrix} n & \mathbf{1}'\boldsymbol{\mu}_0 \\ \boldsymbol{\mu}_0'\mathbf{1} & \boldsymbol{\mu}_0'\boldsymbol{\mu}_0 \end{pmatrix} \begin{pmatrix} \hat{a} - a \\ \hat{b} - b \end{pmatrix} \le F_{2,u}(1-\alpha) \right\} = 1 - \alpha, \tag{15}$$

where $F_{2,u}(1-\alpha)$ is the $(1-\alpha)$-quantile, $\alpha \in (0, 1)$ of the Fisher-Snedecor F-distribution with 2 and u degrees of freedom. By applying Scheffé's Theorem we directly get the

$100 \cdot (1-\alpha)\%$-confidence region for the calibration line $a + b\mu$, for all $\mu \in \langle \mu_l, \mu_u \rangle$:

$$\Pr\left\{\left|(\hat{a}+\hat{b}\mu)-(a+b\mu)\right| \leq \sqrt{2F_{2,u}(1-\alpha)\frac{b_0^2\hat{\sigma}_x^2+\hat{\sigma}_y^2}{m}\left(\frac{1}{n}+\frac{(\mu-\bar{\mu}_0)^2}{\boldsymbol{\mu}_0'\boldsymbol{\mu}_0-n\bar{\mu}_0^2}\right)}\right\}=1-\alpha\,, \quad (16)$$

where $\bar{\mu}_0 = (\mathbf{1}'\boldsymbol{\mu}_0)/n$.

This could be directly used for the multiple-use linear univariate calibration. We note that the calibration line is usually produced (estimated) externally from the viewpoint of the user of the calibrated device. The user of this calibration line (the particular device, instrument) needs for a series of further determinations the following information:

- the sizes n and m of the calibration experiment,
- the estimates (realizations of the estimators) $\hat{a}_{real}$ and $\hat{b}_{real} = b_0 + \hat{\delta}b_{real}$ from the last iteration of the algorithm,
- b_0 and $\boldsymbol{\mu}_0$ from the last iteration of the algorithm,
- the estimates (realizations of the MINQUE estimator) $\hat{\sigma}^2_{x,real}$, $\hat{\sigma}^2_{y,real}$ together with its covariance matrix $\boldsymbol{W}$ based on values from the last iteration of the algorithm.

4 Multiple-Use Calibration

In this section we shall derive a new approximate multiple-use calibration intervals for a series of future determinations based on the approximate (linearized) calibration model (5) with linear constraints (4). We assume that the calibration experiment was realized and the results of the algorithm are available.

We will assume that the future measurement realized by the less precise measurement device A, say x, is a realization of a random variable X, distributed as $X \sim N(\mu_x, \sigma_x^2)$, where μ_x represents the unobservable true value given in units of the less precise instrument and further $\mu_x \in \langle \mu_l, \mu_u \rangle$.

Based on the observed value x we suggest the estimate and a simple derivation of the confidence interval for $\nu_x = a + b\mu_x$ (the unobservable true value in units given by the more precise measuring device B).

First, we suggest to construct the approximate $(1-\gamma)$-confidence region for the calibration line, for small significance level $\gamma \in (0,1)$, chosen by the user according to (16).

Second, for small significance level $\alpha \in (0,1)$, we suggest to construct the approximate $(1-\alpha)$-confidence interval for μ_x, given x and the estimated value (realization of) $\hat{\sigma}_x^2$. For that we suggest to construct the t-statistic with approximate t_v Student's t distribution, where v represents the effective (estimated) degrees of freedom, i.e.

$$t = \frac{X-\mu_x}{\hat{\sigma}_x} \overset{approx}{\sim} t_v\,, \quad (17)$$

where the degrees of freedom are approximated by

$$v = \frac{2\hat{\sigma}_x^4}{\boldsymbol{W}_{11}}\,, \quad (18)$$

where $\hat{\sigma}_x^2$ is the estimate of σ_x^2 given by the algorithm and $\boldsymbol{W}_{11}$ is its variance given by the first element of the estimated covariance matrix $\boldsymbol{W}$. This leads to the approximate $(1-\alpha)$-confidence interval for unobservable value μ_x

$$\mu_x \in \{x \pm \hat{\sigma}_x t_v(1-\alpha/2)\}\,, \tag{19}$$

where $t_v(1-\alpha/2)$ is the $(1-\alpha/2)$-quantile of the Student's t distribution with v degrees of freedom. Let μ_{xl} and μ_{xu} denote the lower and upper limit of the approximate $(1-\alpha)$-confidence interval for μ_x.

The suggested interval estimator for $\nu_x = \langle \nu_{xl}, \nu_{xu} \rangle$ is then given as the intersection of the bounds of the $(1-\gamma)$-confidence region for the whole calibration line $a + b\mu$ and the limits of the $(1-\alpha)$-confidence interval $\langle \mu_{xl}, \mu_{xu} \rangle$ for μ_x. In fact,

$$\begin{aligned}
\nu_{xl} &= \hat{a} + \hat{b}\mu_{xl} - \sqrt{2F_{2,u}(1-\gamma)\frac{b_0^2\hat{\sigma}_x^2 + \hat{\sigma}_y^2}{m}\left(\frac{1}{n} + \frac{(\mu_{xl}-\bar{\mu}_0)^2}{\boldsymbol{\mu}_0'\boldsymbol{\mu}_0 - n\bar{\mu}_0^2}\right)}\,, \\
\nu_{xu} &= \hat{a} + \hat{b}\mu_{xu} + \sqrt{2F_{2,u}(1-\gamma)\frac{b_0^2\hat{\sigma}_x^2 + \hat{\sigma}_y^2}{m}\left(\frac{1}{n} + \frac{(\mu_{xu}-\bar{\mu}_0)^2}{\boldsymbol{\mu}_0'\boldsymbol{\mu}_0 - n\bar{\mu}_0^2}\right)}\,.
\end{aligned} \tag{20}$$

The above $(1-\gamma, 1-\alpha)$-interval estimator for ν_x was suggested in such a way that with probability $(1-\gamma)$ the interval $\langle \nu_{xl}, \nu_{xu} \rangle$ will cover any future (unobservable) value of ν_x with probability approximately equal to $(1-\alpha)$. Our simulation study based on different calibration experiments shows that the the confidence interval (19) has the average empirical confidence level equal to its nominal level $(1-\alpha)$. In particular, we have estimated $\hat{\sigma}_x$ and v by the Algorithm 1 for each generated calibration experiment and then we have calculated the following value of coverage factor (empirical confidence level) which is the same for all possible confidence intervals (19) for given calibration experiment:

$$P = 1 - 2\left(1 - \Phi\left(\frac{\hat{\sigma}_x}{\sigma_x} t_v(1-\alpha/2)\right)\right)\,. \tag{21}$$

The probability $(1-\gamma)$ is the probability that the calibration experiment detected the correct calibration line. Once the calibration line was correctly identified, the probability that any future true value of ν_x will be covered by the interval estimator (20) is approximately (in the above mentioned sense) equal to $(1-\alpha)$.

The preliminary simulations suggest that the empirical coverage probability, i.e. the situation when the true value $\nu_x \in \langle a + b\mu_l, a + b\mu_u \rangle$ is covered by the interval $\langle \nu_{xl}, \nu_{xu} \rangle$, is typically higher than the nominal value $(1-\alpha)$.

This is in contrast with the properties (empirical coverage probabilities) of the single-use calibration confidence interval which was proposed by Wimmer and Witkovský (2005). In this case the empirical coverage probabilities are very close to the nominal level $(1-\alpha)$. Single-use calibration assumes that the independent calibration experiment is conducted for each single future measurement x, the realization of $X \sim N(\mu_x, \sigma_x^2)$. In this case the

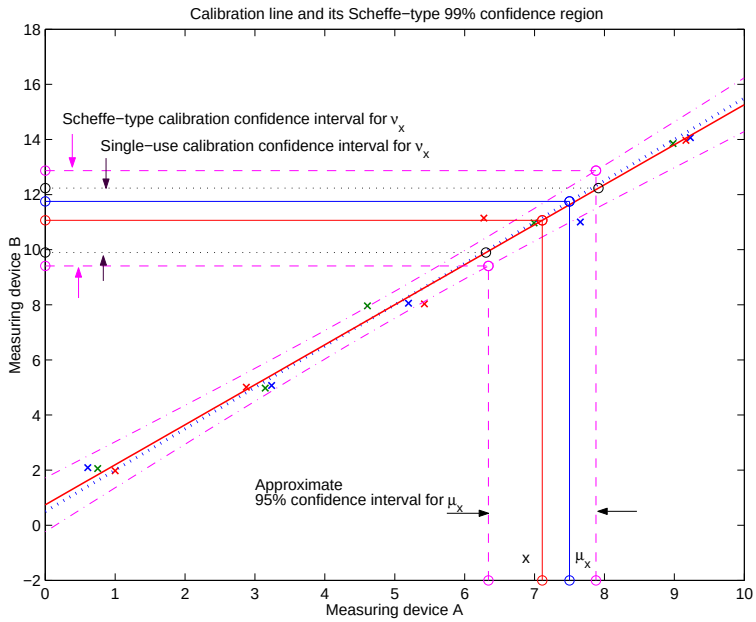

Figure 1: Calibration line and its Scheffé-type confidence region. The thick dotted line represents the true calibration line, the solid line is the estimated calibration line together with the limits of the 99% confidence region (dashed-dotted lines). The dashed lines represent the Scheffé-type interval estimator for $\nu_x = a + b\mu_x$, where $\mu_x = 7.5$, based on $x = 7.1097$, the realization of random variable $X \sim N(7.5, 0.15)$. The dotted lines represent the limits of the approximate 95% single-use calibration confidence interval.

$(1-\alpha)$-confidence interval for ν_x is given by

$$\nu_{xl} = \hat{a} + \hat{b}x - t_u\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{b_0^2\hat{\sigma}_x^2 + \hat{\sigma}_y^2}{m}\left(\frac{1}{n} + \frac{(x-\bar{\mu}_0)^2}{\boldsymbol{\mu}_0'\boldsymbol{\mu}_0 - n\bar{\mu}_0^2}\right) + \hat{b}^2\hat{\sigma}_x^2}\,,$$

$$\nu_{xu} = \hat{a} + \hat{b}x + t_u\left(1-\frac{\alpha}{2}\right)\sqrt{\frac{b_0^2\hat{\sigma}_x^2 + \hat{\sigma}_y^2}{m}\left(\frac{1}{n} + \frac{(x-\bar{\mu}_0)^2}{\boldsymbol{\mu}_0'\boldsymbol{\mu}_0 - n\bar{\mu}_0^2}\right) + \hat{b}^2\hat{\sigma}_x^2}\,, \qquad (22)$$

where u is given by (13). For more details see Wimmer and Witkovský (2005).

4.1 Numerical Example

In order to illustrate the numerical results we illustrate the method for constructing the Scheffé-type calibration confidence interval by the following artificial data. Let $\boldsymbol{\mu} = (1,3,5,7,9)'$ be the vector of true values of $n = 5$ well prepared quantities (primary standards), given in units of the less precise device A. Let $\boldsymbol{X} \sim N(\boldsymbol{\mu}, \sigma_x^2)$ represent the random vector of measurements taken by the instrument A, with variance $\sigma_x^2 = 0.15$. We will repeat the measurements independently $m = 3$ times.

Similarly, let $\boldsymbol{\nu}$ represent the true values of the standards, given in units of the more precise device B, and let $\boldsymbol{\nu} = a + b\boldsymbol{\mu} = 0.5 + 1.5\boldsymbol{\mu}$ be the true calibration line. Let $\boldsymbol{Y} \sim N(\boldsymbol{\mu}, \sigma_y^2)$ represent the random vector of measurements taken by the instrument B,

with $\sigma_y^2 = 0.01$. We will repeat the measurements independently m times. The observed measurements taken during the calibration experiment are given in the table below

$\boldsymbol{\mu}$		$\boldsymbol{x}$			$\boldsymbol{y}$		$\boldsymbol{\nu}$
1	0.6086	0.7507	1.0000	2.0896	2.0569	1.9766	2
3	3.2380	3.1473	2.8769	5.0731	4.9744	5.0118	5
5	5.1966	4.6092	5.4241	8.0578	7.9623	8.0315	8
7	7.6555	6.9924	6.2742	11.0040	10.9704	11.1444	11
9	9.2290	8.9813	9.1658	14.0677	13.8525	13.9649	14

The following estimates (the result of the pre-planned calibration experiment) are given after 10 iterations of the algorithm: $\hat{a} = 0.7405$, $\hat{b} = b_0 = 1.4522$, $\hat{\sigma}_x^2 = 0.1264$, $\hat{\sigma}_y^2 = 0.0057$, $\boldsymbol{\mu}_0 = (0.8933, 2.9497, 5.0123, 7.0897, 9.1048)'$, and $\boldsymbol{W}_{11} = 2.4817 \cdot 10^{-3}$, $\boldsymbol{W}_{22} = 6.4042 \cdot 10^{-6}$, $\boldsymbol{W}_{12} = -6.7842 \cdot 10^{-7}$.

Further, let $x = 7.1097$ be a realization of $X \sim N(\mu_x, \sigma_x^2)$, i.e. the future measurement taken by the less precise instrument A, where $\mu_x = 7.5$ and $\sigma_x^2 = 0.15$, i.e. the true value given in units of the more precise instrument B is $\nu_x = 0.5 + 1.5 \cdot 7.5 = 11.75$. For $\alpha = 0.05$ and by using (19) we get the 95%-confidence interval for μ_x:

$$\mu_x \in \{x \pm \hat{\sigma}_x t_v(1 - \alpha/2)\} = \{7.1097 \pm 0.7688\}\,,$$

with $v = 12.9$ as given in (18).

For $\gamma = 0.01$ and by using the estimated degrees of freedom $u = 13.4$ in (13) we calculate the Scheffé-type $(99\%, 95\%)$-interval estimate for ν_x according to (20):

$$\nu_x \in \langle 9.4096, 12.8699 \rangle\,.$$

For comparison we present also the single-use 95%-confidence interval for ν_x calculated according to (22):

$$\nu_x \in \langle 9.8971, 12.2330 \rangle\,.$$

5 Conclusions

In the paper we have considered the comparative calibration problem. We have suggested a method for construction of the Scheffé-type confidence region for the true (unobservable) calibration line based on the replicated linearized calibration model and on the Kenward-Roger type approximation to the distribution of the F-statistic for the parameters of the calibration line. Further, the interval estimate for the true value of the measurand (in units of the more precise instrument) was suggested in the multiple-use calibration case. The calculation of the interval estimate requires estimates of the parameters of the pre-planned calibration experiment given by the proposed algorithm.

Acknowledgement

The research on the subject was supported by VEGA, the Scientific Grant Agency of Slovak Republic, grants No. 1/0264/03 and 2/4026/04.

References

Brown, P. J. (1993). *Measurement, Regression, and Calibration*. Oxford: Clarendon Press.

Eisenhart, C. (1939). The interpretation of certain regression methods and their use in biological and industrial reserach. *Annals of Mathematical Statistics*, *10*, 162-186.

Kenward, M. G., and Roger, J. H. (1997). Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*, *53*, 983-997.

Krutchkoff, R. G. (1967). Classical and inverse methods of calibration. *Technometrics*, *9*, 425-439.

Kubáček, L., and Kubáčková, L. (2000). *Statistics and Metrology. (In Czech)*. Olomouc: Univerzita Palackého v Olomouci.

Scheffé, H. (1973). A statistical theory of calibration. *The Annals of Statistics*, *1*, 1-37.

Wimmer, G., and Witkovský, V. (2005). Univariate linear calibration via error in variables model. In *Preprints Series, 16/2005.* Mathematical Institute of the Slovak Academy of Sciences, Bratislava, http://www.mat.savba.sk.

Wimmer, G., Witkovský, V., and Savin, A. (2004). Confidence region for parameters in replicated errors in variables model. In J. Antoch (Ed.), *COMPSTAT. Proceedings in Computational Statistics. 16th Symposium Held in Prague* (p. 1987-1994). Heidelberg: Physica-Verlag.

Authors' addresses:

Gejza Wimmer
Masaryk University
Janáčkovo nám. 2a
662 95 Brno, Czech Republic
and
Institute of Mathematics and Computer Science
Banská Bystrica, Slovak Republic
E-mail: wimmer@mat.savba.sk

Viktor Witkovský
Institute of Measurement Science
Slovak Academy of Sciences
Dúbravská cesta 9
841 04 Bratislava, Slovak Republic
E-mail: witkovsky@savba.sk