

# Determination of pressure balance distortion coefficient and zero-pressure effective area uncertainties

V. Ramnath\*

Pressure & Vacuum Laboratory, National Metrology Institute of South Africa, Private Bag X34, Lynnwood Ridge, 0040 Pretoria, South Africa

Received: 23 September 2011 / Accepted: 7 November 2011

**Abstract.** The behaviour of piston-cylinder operated pressure balances is characterized by the distortion coefficient  $\lambda$  and zero-pressure effective area  $A_0$  which model the variation of a pressure balance's area in terms of the applied pressure. This paper determines the uncertainties in  $\lambda$  and  $A_0$  when utilizing the method of cross-floating with another pressure balance standard whose parameters and associated uncertainties are known. A limitation that is frequently encountered in many attempts of the uncertainty analysis for a pressure balance is that no readily accessible uncertainty quantification framework for the distortion coefficient is present. As a result the uncertainty in a pressure balance's area at elevated applied pressures is typically underestimated in the absence of this uncertainty information. We firstly review the uncertainty formulation for a pressure balance generated pressure involving correlation effects in terms of an implicit multivariate matrix equation approach and then utilizing the resulting solution present the methodology to consistently perform the uncertainty analysis for  $\lambda$  and  $A_0$ .

**Keywords:** Pressure balance; effective area; distortion coefficient; uncertainty analysis; parameter estimation

## 1 Introduction

Piston-cylinder operated pressure balances are commonly utilized in laboratories for both primary scientific and secondary industrial pressure scale realizations, and in order to fully achieve the benefit of high accuracy laboratory standards the parameters that are associated in the characterization of these instruments must be quantified. The most general case for the modelling of a pressure balance area  $A$  is in terms of an underlying model in terms of the pressure  $p$  and calibration parameters  $\mathbf{a} = \{a_1, \dots, a_q\}$  represented by an equation  $A(p, \mathbf{a})$ . Different parametric forms of modelling  $A(p, \mathbf{a})$  are possible, but typically polynomial approximations that are either linear or quadratic such as  $A = A_0(1 + \lambda_1 p + \lambda_2 p^2)$  in the case of high pressures as discussed by Sutton et al. [1] or equivalent modified forms in the case of lower pressures such as  $A = A_0(1 + \lambda(p - p_{min})^b)$  where a more pronounced variation in area at lower pressures is apparent that can not be accurately modelled by a straight line or low order polynomial as discussed in Blagojevic and Bajsic [2] are utilized.

In order to realize a pressure scale in practice one must specify the form of  $A(p, \mathbf{a})$  since the pressure is defined implicitly in terms of  $A(p, \mathbf{a})$  for which a general measurand formulation that aligns with the mechanical definition of

pressure  $p = F/A$  where  $F$  is a generalized applied force and  $A$  is a projected area is of form

$$p = \frac{(m + c)g}{A(p, \mathbf{a})[1 + \phi(t)]} \quad (1)$$

for a temperature  $t$  and applied mass  $m$  that is corrected for air buoyancy. In this general formulation  $\phi$  is a temperature compensation function and  $c$  is a measured constant characteristic of the particular pressure balance that are specified for the particular form of the defining equation used to model the pressure balance's area.

Based on linear elasticity theory the choice of the form of equation to model the effect of pressure on area and parameters used in the characterization is for the most part that of a two parameter model  $\mathbf{a} = \{a_1, a_2\}$  incorporating the distortion coefficient  $\lambda$  and zero-pressure area  $A_0$  in the case of pressure balances operated at high pressures, where  $\mathbf{a} = \{A_0, \lambda\}$ . In this paper the variation of the area  $A$  with applied pressure  $P$  is formulated as

$$A = A_0(1 + \lambda P) \quad (2)$$

where the applied pressure  $P$  is defined as  $P = p - p_a$ ,  $p$  is the absolute pressure and  $p_a$  is the atmospheric pressure as per the notation in Dadson et al. [3]. Using this formulation the applied pressure is expressed as  $P = F/A$  in which  $F$  is a generalized force term that accounts for all inherent force components when the the pressure balance

\* Correspondence: [vramnath@nmisa.org](mailto:vramnath@nmisa.org)

is in operation and which can be accurately characterized in terms of the experimental components used to generate the force loading condition.

As a result the uncertainty in area, once the pressure is numerically solved for from the defining modelled pressure balance equation, is completely encapsulated in terms of the underlying parameters of the zero-pressure area  $A_0$  and distortion coefficient  $\lambda$ .

## 2 Review of existing methods for pressure measurand formulation

In general the two predominate traditional methods of solving for the generated pressure are the  $P$  and  $\Delta P$  methods as discussed in Forbes and Harris [4] which utilizing the general formulation for two pressure balances when cross-floated and at equilibrium i.e. at the same applied pressure are

$$\frac{(m+c)g}{A(p, \mathbf{a})[1+\phi(t)]} = \frac{(M+C)g}{A(p, \mathbf{b})[1+\Phi(T)]} \quad (3)$$

where the upper case symbols represent a piston-cylinder 'B' with a known parameter  $\mathbf{b}$  which may be a laboratory standard (LS) and the lower case symbols represent a piston-cylinder 'A' with an unknown parameter  $\mathbf{a}$  such as a transfer standard (TS) or unit under test (UUT). The  $P$ -method involves the optimal solution for the unknown parameter  $\mathbf{a}$  from the equivalent equation

$$\frac{a_1 + a_2 p_i}{b_1 + b_2 p_i} = \frac{(m_i + c)[1 + \Phi(T_i)]}{(M_i + C)[1 + \phi(t_i)]} + e_i \quad (4)$$

where  $e_i$  is an error term at the particular pressure  $p_i$  for a sequence of  $n_P$  pressures i.e.  $i = 1, \dots, n_P$ , whilst the  $\Delta P$ -method involves the elimination of the unknown constant  $c$  through the solution of the equation

$$a_1 + a_2(p_j + p_1) = \frac{(m_j - m_1)g}{\{p_j - p_1 + [\phi(t_j) - \phi(t_1)]p_1\}[1 + \phi(t_j)]} + \varepsilon_j \quad (5)$$

where  $\varepsilon_j$  is an error term for the pressure  $p_j$  for a sequence of  $n_{\Delta P}$  pressures i.e.  $j = 1, \dots, n_{\Delta P}$ . Forbes and Harris [4] performed studies to compare the two methods and determined that the  $P$  method offers superior results in the uncertainty of the zero-pressure effective area  $A_0$  when the mass terms are accurately known.

The  $\Delta P$  method on the other hand is found to exhibit a larger statistical spread in uncertainty for  $A_0$  and also tends to introduce systematic errors if the estimate for the lowest pressure  $p_1$  is incorrect. In general the  $\Delta P$  method is utilized as the preferential method when the mass terms are not accurately known.

More recently the residual mass method developed by Sutton et al. [1] has been proposed which utilizes a generalized least squares (GLS) approach in its method of analysis of the underlying cross-float data starting from a

basis equation of form  $p = (M_e + M)g/A_0c[1 + f(p)] + p_H$  where  $f(p)$  is the equation to model the elastic distortion which in our choice of formulation is  $f(p) = \lambda p$  and where  $M$  is the buoyancy corrected total mass of form  $M = \sum_j m_j(1 - \frac{\rho_a}{\rho_j})$ . The least squares analysis is based on the insight that the cross-floated pressure data can be formulated in the form

$$M_i = M_e(-1) + A_0 \left( \frac{c_i p_i}{g} \right) + (\lambda A_0) \left( \frac{c_i p_i^2}{g} \right)^2 \quad (6)$$

which is amenable to a linear least squares analysis for the solution of the unknowns  $M_e$ ,  $A_0$  and  $\lambda A_0$  of a transfer standard in the case of a known generated pressure standard. Extensions to cases of two reference pressure balances i.e. a laboratory standard and transfer standard and for more complex elastic distortion models e.g.  $f(p) = \lambda_1 p + \lambda_2 p^2$  are possible in which case additional unknown terms to solve for result from the linear least squares formulation  $\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e}$  where in the present instance  $\mathbf{y}$  is a  $q \times 1$  vector of measurement results i.e.  $y_i$  is the total experimental loaded mass  $M_i$  corresponding to generated pressure  $p_i$ ,  $\mathbf{X}$  is a  $q \times h$  design matrix where the  $i$ th row is  $(-1, c_i p_i/g, c_i p_i^2/g)$ ,  $\boldsymbol{\beta}$  is a  $h \times 1$  vector of the unknowns to be solved for i.e.  $\boldsymbol{\beta} = (M_e, A_0, \lambda A_0)$  in the case for  $h = 3$ , and  $\mathbf{e}$  is a  $q \times 1$  vector of random disturbances.

We remark that in this analysis that the distortion coefficient  $\lambda$  is inferred as  $\lambda = \beta_3/\beta_2$  where  $\beta_1 = M_e$ ,  $\beta_2 = A_0$  and  $\beta_3 = \lambda A_0$  are the solved parameters and that there is no directly amenable calculation for the uncertainty in  $\lambda$  since the components in the least squares parameter  $\boldsymbol{\beta} = \{\beta_1, \beta_2, \beta_3\}$  do not directly correspond to the pressure balance's area measurand parameter  $\mathbf{a} = \{a_1, a_2\}$ . The inherent complication in this analysis is the presence of product terms of non-independent random variables which necessitates the knowledge of covariances, of which the statistical calculations as discussed in detail by Wilkinson [5], is required to estimate a measurand  $y = f(x_1, \dots, x_n)$  uncertainty  $u(y)$  from

$$u^2(y) = \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + \sum_{i=1}^n \sum_{j=1}^n \left\{ \frac{1}{2} \left( \frac{\partial^2 f}{\partial x_i \partial x_j} \right)^2 + \frac{\partial f}{\partial x_i} \frac{\partial^3 f}{\partial x_i \partial x_j^2} \right\} \times u^2(x_i) u^2(x_j) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j) \quad (7)$$

according to the GUM [6] where the correlation coefficient is related to the covariance according to the formula

$$r(x_i, x_j) = \frac{u(x_i, x_j)}{u(x_i) u(x_j)}, \quad -1 \leq r(x_i, x_j) \leq 1 \quad \forall i, j \in [1, \dots, n]. \quad (8)$$

The GLS solution for the estimate of the unknown from standard statistical analysis [1] is  $\hat{\beta} = \hat{C} \mathbf{X}^T \hat{\Phi}^{-1} \mathbf{y}$  where  $\hat{\Phi}$  is the symmetric uncertainty matrix where the diagonal elements equal the variances  $u^2(y_i)$  i.e.  $u^2(M_i)$  of the measurement results and  $\hat{C} = (\mathbf{X}^T \hat{\Phi}^{-1} \mathbf{X})^{-1}$  is a uncertainty matrix, where the diagonal elements of  $\hat{C}$  correspond to the variances of  $\hat{\beta}$ .

An example of one limitation of the residual mass method occurs when comparing similar terms in the residual mass method and the full expression cited in Dadson et al. [3] that through a comparison of force terms it is seen that, ignoring a height term for hydrostatic pressure correction, the resultant force must equal the weight of the form  $[M - V_s(\rho_f - \rho_a) + \frac{\sigma C}{g}]g = (M + M_e)g$  or equivalently  $\frac{\sigma C}{g} - V_s(\rho_f - \rho_a) = M_e$ . The limitation within this assumption is that the residual mass term  $M_e$  is assumed to be constant across the entire pressure range and so although the buoyancy correction term of the submerged piston will vary with pressure e.g. through the use of an ideal gas equation of state or alternately the application of a functional for modelling the working fluid's variation in density such as a silicon oil with varying pressure there is no mechanism to refine or incorporate such physical characteristics within the GLS framework which is an implicit compromise within the adoption of a linear least squares analysis.

A further limitation that occurs in a GLS approach in the form that is amenable to a linear least squares analysis is that the variances in  $\hat{C}$  relate only to the uncertainties included in  $\hat{\Phi}$  and as such only uncertainties in the experimental mass loads can be explicitly controlled since other uncertainty contributions in terms of the covariances i.e. off-diagonal elements in the matrix  $\hat{\Phi}$  are not easily quantified and other uncertainty contributing factors must be grouped within the residual mass  $M_e$  term. Whilst a method exists as proposed in the original residual mass method by Sutton et al. [1] to incorporate independent uncertainty quantity contributions by propagating these effects through repeated GLS analyses in a manner reminiscent to statistical jackknifing [7] and computing an analogous sensitivity coefficient the mechanism does not allow for the incorporation of correlation effects in the underlying model inputs.

The residual mass method is nevertheless an improvement on the original  $P$  and  $\Delta P$  methods in that it solves for the piston-cylinder parameter  $\mathbf{a}$  of the zero-pressure area and distortion coefficient along with a residual mass  $M_e$  i.e.  $\mathbf{a} = \{A_0, \lambda, M_e\}$  that may be considered as an additional control parameter to refine the model parameter fitting in the case of a combination of measurement data that exhibits slight experimental errors and a measurand model that incorporates working approximations.

As a result the residual mass  $M_e$  may be employed to account for any unweighed parts of the piston and loading components such as surface tension terms which are not accurately accounted for as a means to ensure that the underlying defining generated pressure equations

are consistent without introducing an excessive associated uncertainty in the computed generated pressure. Alternatives to linear least squares limitations and incorporation of possible nonlinear and correlation effects for confidence limit estimation on the model parameters include the well known Monte Carlo techniques such as the boot strap method as originally discussed in Press et al. [8] and later by Cox and Siebert [9] from a general implementation perspective, and more recently by Ramnath [10] who utilized and fully developed these alternatives for the computation of the effective area uncertainty determination of apex piston-cylinder based pressure standards by a direct Monte Carlo numerical simulation for a primary scientific pressure standard characterization.

Previous work performed by Zvizdic et al. [11] utilizing the  $P$ -method has outlined a simplified method for the uncertainty estimation of the effective area for hydraulic piston-cylinder pressure balances operated in gauge mode. This work utilized standard statistical linear regression formulae to provide estimates for the zero-pressure area  $A_0$  and gradient of the best fit straight line, and developed uncertainty estimates for the generated pressure and zero-pressure area  $A_0$  but did not fully develop a ready means for the estimation of uncertainty of the distortion coefficient  $\lambda$ .

In this paper our primary approach is to adopt the  $P$  method as the starting point in the analysis in order to develop the methodology for computing the uncertainties in  $A_0$  and  $\lambda$ . We then elaborate on an approximate method to incorporate additional model parameters such as the residual mass term into a nonlinear least squares solution along with its uncertainty estimate and the effect this has in the final computed uncertainties for  $A_0$  and  $\lambda$ .

### 3 Mathematical model for pressure balance

The measurand model that we will consider in this paper taking into account all force components is of form:

$$p - p_a = \frac{1}{S} \left( \left\{ \sum_{i=1}^n m_i \left( 1 - \frac{\rho_a}{\rho_i} \right) - V_s (\rho_f - \rho_a) + H (\rho_f - \rho_a) S \right\} g + \sigma C \right) \quad (9)$$

$$S(t) = A_0 [1 + \lambda P] \phi(t, t_{ref}) \quad (10)$$

where  $p_a$ ,  $t_a$  and  $h_a$  represents the atmospheric pressure, temperature and relative humidity respectively following the notation of Picard et al. [12] where the relative humidity is represented by a number such that  $0 \leq h_a \leq 1$ .

In the above model  $S$  is the actual instantaneous area at temperature  $t$  and by setting  $\alpha = \alpha_p + \alpha_c$  where  $\alpha_p$  and  $\alpha_c$  are the linear thermal expansion coefficients of the piston and cylinder respectively one may account for the thermal expansion/contraction of the piston-cylinder when there is a temperature difference  $(t - t_{ref})$  for  $t \neq t_{ref}$  by constructing a temperature compensation  $\phi(t, t_{ref})$  as

$$\phi(t, t_{ref}) = 1 + \alpha (t - t_{ref}) \quad (11)$$

where in practice one commonly sets  $t_{ref}$  to 20°C or occasionally to 23°C. In order to utilize the temperature compensation function one must set the reference temperature to the temperature at which the laboratory standard piston-cylinder zero-pressure area  $A_0$  was either dimensionally characterized or alternately experimentally determined via a cross-float with another standard.

In the measurand model  $n$  is the number of mass pieces used,  $m_i$  is the true mass of the  $i$ th mass piece where one may convert [13] from a conventional mass  $m_c$  typically supplied from a mass laboratory calibration report to a true mass  $m$  using the formula  $m = m_c(1 - 1.2/8000)(1 - 1.2/\rho)$  where  $\rho$  is the actual density e.g. 7920 kg m<sup>-3</sup>,  $\rho_i$  is the density of the  $i$ th mass piece,  $g$  is the local gravitational acceleration constant,  $\rho_a$  is the density of the atmosphere to which the mass pieces are exposed,  $V_s$  is the submerged volume of the piston within the working fluid,  $\rho_f$  is the density of the working fluid,  $\gamma$  is the surface tension of the working fluid, and  $C$  is the circumference of the piston in contact with the working fluid that may be simply approximated using  $S = \frac{\pi}{4}D^2$  and  $C = \pi D$  as

$$C = \sqrt{4\pi S} \quad (12)$$

where the surface tension  $\gamma$  is only present in the case of a fluid-gas interface [14]. Depending on the whether the working fluid is a liquid or gas different defining equations can be utilized to formulate the fluid density  $\rho_f$  in terms of the unknown pressure  $p$ . For gases where the continuum assumption is valid at non-vacuum pressures the equation of state  $p = \rho RT$  for an ideal fluid is a convenient choice, with  $R = \mathbf{R}/M$  where  $\mathbf{R}$  is the universal gas constant [15] and  $M$  is the gas molecular weight. For liquids such as oils which are used at high pressures the formulation for fluid density  $\rho_f$  of form

$$\rho_f = \left[ 1 - \frac{p - p_{f0}}{E_f} \right]^{-1} \left[ \frac{\rho_{f0}}{1 + \beta(t_f - t_{f0})} \right] \quad (13)$$

where  $\beta$  is the fluid volumetric thermal expansion coefficient,  $\rho_{f0}$ ,  $t_{f0}$  and  $p_{f0}$  the reference density, temperature and pressure at the initial state,  $E$  the fluid bulk modulus elasticity defined as  $E = -V\partial p/\partial V$ , where  $V$  is the volume of the fluid, represents a general purpose approximation in the absence of further information that can yield sufficiently accurate results for moderate pressure ranges.

In the the illustrative example that follows we will assume that the pressure balance uses di(2-ethylhexyl) sebecate ('DHS') as the working fluid for which the equation

$$\rho_f = \left[ 912.7 + 0.752(p/10^6) - 1.65 \times 10^{-3}(p/10^6)^2 + 1.5 \times 10^{-6}(p/10^6)^3 \right] \left[ 1 - 7.8 \times 10^{-4}(t_f - 20) \right] \quad (14)$$

is known from the literature to accurately model the density dependance on pressure and temperature [16].

If the pressure balance is operated in absolute mode then the atmospheric pressure  $p_a$  above the piston will typically consist of a working gas such as nitrogen in which case the density  $\rho_a$  can be calculated from the equation of state, whilst if the pressure balance is operated in gauge mode then  $\rho_a$  can be calculated in terms of  $p_a$ ,  $t_a$  and  $h_a$  as per the CIPM-2007 formula for the density of moist air [12].

By rearranging the defining generated pressure measurand equation we obtain a nonlinear equation  $f(Y|\mathbf{X}) = 0$  as

$$f = \sum_{i=1}^n m_i \left( 1 - \frac{\rho_a}{\rho_i} \right) - V_s(\rho_f - \rho_a) + \frac{\sigma C}{g} + H(\rho_f - \rho_a)A_0[1 + \lambda(p - p_a)]\phi(t, t_{ref}) - \frac{1}{g}(p - p_a)A_0[1 + \lambda(p - p_a)]\phi(t, t_{ref}) \quad (15)$$

where

$$\mathbf{X} = [m_1, \dots, m_n, \rho_1, \dots, \rho_n, A_0, \lambda, \alpha, t, V_s, H, p_a, t_a, h_a, \sigma, g]^T \quad (16)$$

$$= [x_1, \dots, x_{2n+11}]^T \quad (17)$$

represents a multi-dimensional known input parameter and  $Y$  representing the pressure is the unknown root to be solved for, noting that in the most general case that certain terms such as the fluid density or circumference may be to varying degrees of refinement/complexity be defined in terms of the pressure. In the above model it is common to make the simplifying assumption that the pressure balance piston-cylinder is in thermodynamic equilibrium with the working fluid i.e.  $t_f = t$  after a suitable stabilization time although this is not essential and one may if necessary model the heating effect on the working fluid as the pressure is increased.

For the most general case using the formula for  $f$  from the multivariate implicit model following the approach of Cox and Harris [17] results in a vector-matrix equation of form

$$\mathbf{h}(\mathbf{Y}, \mathbf{X}) = \mathbf{0} \quad (18)$$

where  $\mathbf{h}(\mathbf{Y}, \mathbf{X}) = [f_1, \dots, f_m]^T$  now represents a system of nonlinear equations and  $f_i$  corresponds to the generated pressure  $p_i$  for  $q$  generated pressures in the range  $[p_{min}, p_{max}]$ . The numerical solution  $\mathbf{y}^{(0)} = [p_1, \dots, p_q]^T$  obtained is then used as an input into the matrix equation

$$\mathbf{J}_y \mathbf{V}_y \mathbf{J}_y^T = \mathbf{J}_x \mathbf{V}_x \mathbf{J}_x^T \quad (19)$$

to solve for the uncertainty matrix  $\mathbf{V}_y$  where the diagonal elements correspond to the variances  $V_{ii} = \text{Var}(y_i)$  of the generated pressures  $p_1, \dots, p_q$  and the off-diagonal elements correspond to the covariances  $V_{ij} = \text{Cov}(y_i, y_j)$  with  $1 \leq i, j \leq m$  and  $i \neq j$ . In the simpler case where we do not model the correlation effect of the input parameter  $\mathbf{X}$  on the sequence of generated pressures  $p_1, \dots, p_q$  but merely the covariance matrix  $\mathbf{V}_x$  we have instead

$$u^2(y) \left( \frac{\partial f}{\partial y} \right)^2 = (\nabla_{\mathbf{x}} f) \mathbf{V}_x (\nabla_{\mathbf{x}} f)^T \quad (20)$$

which yields the uncertainty in generated pressure  $u(p)$  where  $\nabla_{\mathbf{x}} f = [\partial f/\partial x_1, \dots, \partial f/\partial x_{(2n+1)}]^\top$ . In specifying the input  $\mathbf{V}_{\mathbf{x}}$  we assume no correlation effects of the input parameter  $\mathbf{X}$  for simplicity and remark that the only possible significant non-zero correlation effect that may occur is between the  $A_0$  and  $\lambda$  of the laboratory standard used.

We remark that generally there will always be a high correlation between  $A_0$  and  $\lambda$  in the event that  $A_0$  and  $\lambda$  are obtained from the cross-floating of another precision piston-cylinder standard, however in most instances there may be no accessible information on the covariance between these two quantities and as a result equation (20) is generally more feasible to implement.

#### 4 Implementation of method to determine cross-floated area

Utilizing the  $P$  method it follows that at equilibrium for two piston-cylinders, with ‘A’ signifying the piston-cylinder of known characteristics i.e. the laboratory standard and ‘B’ signifying the piston-cylinder with unknown characteristics to be determined i.e. the unit-under-test (UUT) that the basis equation reads as

$$0 = \sum_{i=1}^n m_i \left(1 - \frac{\rho_a}{\rho_i}\right) - V_s (\rho_f - \rho_a) + \frac{\sigma \sqrt{4\pi S}}{g} + H (\rho_f - \rho_a) S - \frac{P}{g} S. \quad (21)$$

By making the assumption that the fluid-gas interface i.e. the circumference  $C = \sqrt{4\pi S}$  of the piston has minimal change with varying temperature and pressure i.e.

$$\frac{\partial S}{\partial t} (\delta t) = A_0 [1 + \lambda (p - p_a)] (\alpha_p + \alpha_c) (\delta t) \ll 1$$

$$\frac{\partial S}{\partial p} (\delta p) = A_0 \lambda [1 + (\alpha_p + \alpha_c) (t - t_0)] (\delta p) \ll 1$$

where  $\delta t = \mathcal{O}(u(t))$  and  $\delta p = \mathcal{O}(u(p))$  are small perturbations in temperature and pressure respectively we have that the piston-cylinder area  $S$  is by setting  $C$  to a constant that

$$S = \{P - H (\rho_f - \rho_a) g\}^{-1} \left[ \sum_{i=1}^n m_i g \left(1 - \frac{\rho_a}{\rho_i}\right) + \sigma C - V_s (\rho_f - \rho_a) g \right]. \quad (22)$$

The height correction term is set as  $H = 0$  since the computed applied pressure  $P = p - p_a$  already includes the height correction adjustment from which the area prior to any temperature corrections is

$$S(t) = \frac{1}{P} \left[ \sum_{i=1}^n m_i g \left(1 - \frac{\rho_a}{\rho_i}\right) + \sigma C - V_s (\rho_f - \rho_a) g \right]. \quad (23)$$

As previously discussed in general access to information on covariances is either not available or difficult to directly estimate using readily available statistical routines so the use of the more general formulae for the calculation of a sequence of pressures is avoided, and the formulae for a single pressure point of a cross-float is used for a set of pressure points within a range of pressures.

In order to implement a cross-float calibration using a Laboratory Standard (LS) viz. a piston-cylinder ‘A’ against a Transfer Standard (TS) viz. a piston-cylinder ‘B’ we follow the following steps for a sequence of  $N$  pressure points as:

- calculate the applied pressure  $P$  of the LS using equation (15) and its corresponding uncertainty using equation (20)
- use the calculated generated applied pressure  $P$  to calculate the corresponding TS area  $S$  with equation (23)
- calculate the area  $A$  adjusted to the reference temperature  $t_0$  using equation (11) as  $A = S/\phi(t, t_{ref})$  along with its corresponding standard uncertainty  $u(A)$ .

The result of this will be a set of applied pressures and corresponding areas of the TS, along with their associated standard uncertainties, which may be conveniently expressed in vector form as

$$\mathbf{x} = [P_1, \dots, P_N]^\top \quad (24)$$

$$\mathbf{u}(\mathbf{x}) = [u(P_1), \dots, u(P_N)]^\top \quad (25)$$

$$\mathbf{y} = [A_1, \dots, A_N]^\top \quad (26)$$

$$\mathbf{u}(\mathbf{y}) = [u(A_1), \dots, u(A_N)]^\top \quad (27)$$

which form the data to be analysed to determine the TS distortion coefficient  $\lambda$  and zero-pressure area  $A_0$ .

The uncertainty in the TS effective area may be determined using the GUM [6] as

$$u^2(A) = (\partial_S A)^2 u^2(S) + (\partial_\alpha A)^2 u^2(\alpha) + (\partial_t A)^2 u^2(t)$$

$$= \left\{ [1 + \alpha (t - t_{ref})]^{-1} \right\}^2 u^2(S)$$

$$+ \left\{ -S [1 + \alpha (t - t_{ref})]^{-2} (t - t_{ref}) \right\}^2 u^2(\alpha)$$

$$+ \left\{ S [1 + \alpha (t - t_{ref})]^{-2} \alpha \right\}^2 u^2(t) \quad (28)$$

$$u^2(S) = \sum_{i=1}^n \left[ (\partial_{m_i} S)^2 u^2(m_i) + (\partial_{\rho_i} S)^2 u^2(\rho_i) \right]$$

$$+ (\partial_P S)^2 u^2(P) + (\partial_{\rho_a} S)^2 u^2(\rho_a)$$

$$+ (\partial_{\rho_f} S)^2 u^2(\rho_f) + (\partial_{V_s} S)^2 u^2(V_s)$$

$$+ (\partial_g S)^2 u^2(g) + (\partial_\sigma S)^2 u^2(\sigma) + (\partial_C S)^2 u^2(C) \quad (29)$$

where

$$\frac{\partial S}{\partial m_i} = \frac{g}{P} \left(1 - \frac{\rho_a}{\rho_i}\right) \quad (30)$$

$$\frac{\partial S}{\partial \rho_i} = \frac{m_i g \rho_a}{P \rho_i^2} \quad (31)$$

$$\frac{\partial S}{\partial P} = \frac{-S}{P} \quad (32)$$

$$\frac{\partial S}{\partial \rho_a} = \sum_{i=1}^n \left[ \frac{-m_i g}{P \rho_i} \right] + \frac{V_s g}{P} \quad (33)$$

$$\frac{\partial S}{\partial \rho_f} = \frac{-V_s g}{P} \quad (34)$$

$$\frac{\partial S}{\partial V_s} = \frac{-(\rho_f - \rho_a) g}{P} \quad (35)$$

$$\frac{\partial S}{\partial g} = \frac{-V_s (\rho_f - \rho_a)}{P} + \sum_{i=1}^n m_i \left(1 - \frac{\rho_a}{\rho_i}\right) \quad (36)$$

$$\frac{\partial S}{\partial \sigma} = \frac{C}{P} \quad (37)$$

$$\frac{\partial S}{\partial C} = \frac{\sigma}{P} \quad (38)$$

noting that

$$u^2(P) = u^2(p) + u^2(p_a) \quad (39)$$

$$u^2(\rho_f) = \left(\frac{\partial \rho_f(p, t)}{\partial p}\right)^2 u^2(p) + \left(\frac{\partial \rho_f(p, t)}{\partial t}\right)^2 u^2(t) \quad (40)$$

using either equation (40) or a similar defining equation in the case of a particular oil used as a working fluid or alternately the equation of state  $p = \rho RT$  when a gas such as dry nitrogen is used. For the majority of cases for high pressure measurements the uncertainty of the air density when utilizing the CIPM-2007 formula may be approximated as

$$u^2(\rho_a) = \left(\frac{\partial \rho_a}{\partial t_a}\right)^2 u^2(t_a) + \left(\frac{\partial \rho_a}{\partial p_a}\right)^2 u^2(p_a) + \left(\frac{\partial \rho_a}{\partial h_a}\right)^2 u^2(h_a) + \varepsilon^2 \quad (41)$$

where  $\varepsilon$  is a term to account for the underlying uncertainty in the CIPM-2007 equation for the air density factoring in consideration of the mole fraction of argon and other terms and which is  $\varepsilon = (22 \times 10^{-6})\rho_a$  according to the literature [12].

## 5 Method implementation for uncertainty analysis of curve fit parameters

In general there are two basic approaches to follow to determine  $\lambda$  and  $A_0$ , namely that of:

- (i) a direct parameter fit that implicitly uses the data;
- (ii) a ratio approach that explicitly links the TS distortion coefficient value to the LS.

In both approaches there is a need to fit a straight line of form  $y = ax + b$  to coordinate data points  $\mathbf{x} = \{x_1, \dots, x_n\}$  and  $\mathbf{y} = \{y_1, \dots, y_n\}$  to determine the parameters  $a$  and  $b$  along with their associated uncertainties. The  $x_i$  inputs to the curve fitting routine are the calculated generated pressures ( $P_i \pm u(P_i)$ ) and the  $y_i$  inputs are the calculated areas ( $A_i \pm u(A_i)$ ) as previously discussed, for which the covariance  $\text{Cov}(x_i, y_i)$  may be approximately estimated using the discrete data as

$$\text{Cov}(x_i, y_i) \approx \frac{1}{r-1} \sum_{i=1}^r (x_i - \bar{x})(y_i - \bar{y}) \quad (42)$$

for  $r$  pressure and area values consistent with equation (23) as per the GUM [6] for the ‘statistical’ covariance, where the pressures may be generated from sampled values with  $P \sim N(\mu, \sigma^2)$  i.e. where the pressure follows a normal distribution with mean  $\mu$  and variance  $\sigma^2$  where the mean corresponds to the  $i$ th generated pressure and the variance to the calculated standard uncertainty squared. It should be noted that this is in fact a working approximation since in general there may be a variation in ‘logical’ and ‘statistical’ correlations as discussed in Siebert et al. [18]. In the absence of further information one may generally assume  $\text{Cov}(P, A) = 0$  without significant detrimental impact.

An algorithm WTLS-C to accomplish the curve fitting in a straightforward manner has recently been published in the literature [19] for which a condensed summary and program listing illustrating the implementation of the technique is provided for in the appendix. For method (i) i.e. the direct parameter fitting in the case of the  $P$ -method it follows from  $y = ax + b \Leftrightarrow A = A_0(1 + \lambda P)$  that with  $A_0 = b$  and  $\lambda = \frac{a}{b}$  when applying the correlated weighted least-squares curve fit that

$$A_0 = b \pm u(b) \quad (43)$$

$$\lambda = \frac{a}{b} \pm \left[ \left(\frac{1}{b}\right)^2 u^2(a) + \left(\frac{-a}{b^2}\right)^2 u^2(b) + 2 \left(\frac{1}{b}\right) \left(\frac{-a}{b}\right) \times \text{Cov}(a, b) \right]^{1/2} \quad (44)$$

where  $u(a) = \sqrt{\text{Var}(a)}$ ,  $u(b) = \sqrt{\text{Var}(b)}$ , and  $\text{Cov}(a, b)$  are the algorithm’s outputs, where the covariance between the fitted parameters  $a$  and  $b$  is in general not necessarily equal to zero even if the covariance between the applied pressure and area is explicitly set to zero in the absence of adequate statistical knowledge.

Although the application of nonzero correlation coefficients  $\rho_k, k = 1, \dots, n$  for  $n$  experimental points e.g.  $n = 10$  for pressures  $P = 50, \dots, 500$  MPa in the WTLS-C will almost always reduce the uncertainties in the fitted straight line parameters as elaborated by Krystek and Anton [19] they caution that this will only lead to better results if the assumed correlations are indeed present and known, else potentially erroneous results will occur. Due to the challenges both on a practical laboratory physical and statistical level it is thus advantageous to simply

assume zero correlation and err on the side of caution when specifying the covariance term as  $\text{Cov}(P, A) = 0$ . This approach is consistent with the principle of maximum entropy encountered in Bayesian probability or ‘information entropy’.

In the context of metrology uncertainty analysis this is essentially restated as the choice of the probability density function (PDF) is that which results in the largest uncertainty within the specified constraints with the result that no additional assumptions or biases are introduced in the calculations [20]. Since the covariance term  $\text{Cov}(a, b)$  as computed with the WTLS-C algorithm implementation is known and verified in the case  $\text{Cov}(a, b) \neq 0$  it is then appropriate and statistically correct to utilize this information knowledge in the reduction and specification of upper bounds of the calculated uncertainties of the curve fit parameters.

Whilst the computation of the covariance of products of non-independant random variables is not easily tractable from first principles or well known an approximate estimate for the covariance may be utilized from a derivation based on the general expression for the covariance of the products of  $k$  random variables developed by Wilkinson [5] in the special case  $k = 2$  i.e. for the product of two random variables as

$$\text{Cov}(x, y) = \frac{1}{2} \left[ \frac{\text{Var}(xy)}{XY} - \frac{y}{x} \text{Var}(x) - \frac{x}{y} \text{Var}(y) \right] \quad (45)$$

and for the variance of a product of independant random variables as

$$\text{Var}(xy) = X^2 \text{Var}(y) + Y^2 \text{Var}(x) + \text{Var}(x) \text{Var}(y). \quad (46)$$

The purpose for the utilization of the covariance is to relate the two sets of underlying parameter uncertainties

$$\begin{array}{c} y = ax + b \\ \left\{ \begin{array}{c} u(a) \\ u(b) \\ \text{Cov}(a, b) \end{array} \right\} \end{array} \quad \begin{array}{c} A = A_0(1 + \lambda P) \\ \left\{ \begin{array}{c} u(A_0) \\ u(\lambda) \\ \text{Cov}(A_0, \lambda) \end{array} \right\} \end{array}$$

with a view in reducing the uncertainties of  $A_0$  and  $\lambda$  through the presence of non-zero correlation since the application of the GUM [6] in calculating uncertainty includes both variances as well as the covariances of the inputs. In the weighted least squares curve fit of applied pressure  $P_i$  and area  $A_i$  input data points for  $n$  pressures even if there is an absence of specified correlation in the underlying data i.e. if one is conservative and sets  $\text{Cov}(P_i, A_i) = 0 \forall i = 1, \dots, n$  the curve fit will still result in a correlation of the fitted parameters  $a$  and  $b$ . As a result it follows that  $A_0 = b$  and  $\lambda = a/b$  from which with the application of the GUM i.e. equation (7) we have the inclusion of covariances in the uncertainties of the parameters  $a$  and  $b$ .

For method (ii) i.e. the ratio method we follow a similar approach however instead of underlying pressure and area data points we calculate the ratio of effective areas at a particular pressure point i.e.  $R = A^{(TS)}/A^{(LS)}$  where  $A^{(TS)}$  is the area for the unknown piston-cylinder

transfer standard ‘B’ and  $A^{(LS)}$  is the area for the known piston-cylinder laboratory standard ‘A’, and both areas have incorporated temperature adjustments to the particular reference temperature.

Following the approach of the  $P$ -method at equilibrium during cross-floating i.e. equal pressures with  $P^{(A)} = P^{(B)}$  we have  $F^{(A)}/S^{(A)}(t^{(A)}) = F^{(B)}/S^{(B)}(t^{(B)})$  so that on rearranging and taking the ratio of areas

$$\frac{F^{(A)}}{F^{(B)}} = \frac{S^{(A)}}{S^{(B)}} = R. \quad (47)$$

Applying the temperature compensation function definition  $A = S(t)\phi(t)$  and eliminating the temperature dependence we then have

$$\frac{F^{(A)}}{F^{(B)}} = \frac{A_0^{(A)} [1 + \lambda^{(A)} P] \frac{1}{\phi(t^{(A)})}}{A_0^{(B)} [1 + \lambda^{(B)} P] \frac{1}{\phi(t^{(B)})}} \quad (48)$$

which on rearranging may be written as

$$\begin{aligned} \frac{F^{(A)}}{F^{(B)}} &= \left( \frac{A_0^{(A)}}{A_0^{(B)}} \right) \left[ 1 + \left( -\lambda^{(B)} + \lambda^{(A)} \right) P \right. \\ &\quad + \left( \left( \lambda^{(B)} \right)^2 - \lambda^{(A)} \lambda^{(B)} \right) P^2 \\ &\quad + \left. \left( \left( -\lambda^{(B)} \right)^3 + \lambda^{(A)} \left( \lambda^{(B)} \right)^2 \right) P^3 + \varepsilon \right] \\ &\quad \times \frac{[1 + \alpha^{(B)} (t^{(B)} - t_0)]}{[1 + \alpha^{(A)} (t^{(A)} - t_0)]} \\ &\approx R_0 \left[ 1 + \left( \lambda^{(A)} - \lambda^{(B)} \right) P \right] \frac{\phi(t^{(B)})}{\phi(t^{(A)})} \quad (49) \end{aligned}$$

where  $\varepsilon$  is an error term due to the truncation of the Taylor series expansion. Dadson et al. [3] using the assumption  $p_a^{(A)} = p_a^{(B)}$  and  $\rho_f^{(A)} = \rho_f^{(B)}$  with additional simplifications for the surface tension and air buoyancy effects in the case of oil operated pressure balances were able to derive the following working approximation for the area ratio as

$$\begin{aligned} R &\approx \left\{ \sum_{i=1}^{n^{(A)}} m_i^{(A)} - V_s^{(A)} (\rho_f - \rho_a) + (H_A - H_B) \right. \\ &\quad \times (\rho_f - \rho_a) S^{(A)} \left. \right\} \frac{[1 + \alpha (t^{(B)} - t^{(A)})]}{\sum_{j=1}^{n^{(B)}} m_j^{(B)} - V_s^{(B)} (\rho_f - \rho_a)} \quad (50) \end{aligned}$$

in the special case that  $\alpha^{(A)} = \alpha^{(B)}$ . For this expression a reference height level must be chosen and the respective heights of piston ‘A’ and ‘B’ consistently specified relative to this reference level using a chosen sign notation [21]. A convenient sign notation to adopt is to specify displacements in the direction of gravity as positive and to check if the term  $(H_A - H_B)(\rho_f - \rho_a)S^{(A)}$  correctly accounts for the

hydrostatic head correction. As an example if piston 'A' is below piston 'B' then it is required that  $(H_A - H_B) > 0$  and visa versa.

A full explicit analytical solution that does not rely on any simplifying assumptions is available when solved for with a computer algebra system (CAS) such as Maxima [22] for which the area ratio is defined in terms of variables

- $m_i, i = 1, \dots, n$  and  $\rho_i, i = 1, \dots, n$ ;
- $\rho_a(t_a, p_a, h_a)$ ;
- $V_s, H, \sigma$  and  $C$ .

common to both the laboratory standard 'A' as well as the transfer standard 'B' along with the gravitational acceleration  $g$  which appears explicitly in both the approximate formula in equation (50) as well as a full explicit formula when solved with a CAS.

As a result the underlying data set is now, whether with the application of the approximate formula or a full explicit CAS solution, the following vectors

$$\mathbf{x} = \{P_1, \dots, P_N\} \quad (51)$$

$$\mathbf{u}(\mathbf{x}) = \{u(P_1), \dots, u(P_N)\} \quad (52)$$

$$\mathbf{z} = \{R_1, \dots, R_N\} \quad (53)$$

$$\mathbf{u}(\mathbf{z}) = \{u(R_1), \dots, u(R_N)\} \quad (54)$$

where  $R = R^{(A)}/R^{(B)}$  is the ratio of the LS area to the TS area with a temperature adjustment. One then utilizes the WTLS-C algorithm implementation to fit a straight line  $z = mx + c$  to the underlying data points  $\mathbf{x}$  and  $\mathbf{z}$  along with the associated uncertainties to determine the parameters  $R_0$  and  $\lambda_{AB}$  according to a equation

$$R = R_0 (1 + \lambda_{AB}P) \quad (55)$$

so that the parameters are

$$R_0 = c \quad (56)$$

$$\lambda_{AB} = \frac{m}{c} \quad (57)$$

from which it follows that

$$\lambda^{(B)} = \left[ \lambda^{(A)} - \frac{m}{c} \right] \pm \left[ u^2(\lambda^{(A)}) + \left(\frac{1}{c}\right)^2 u^2(m) + \left(\frac{m}{c^2}\right)^2 u^2(c) \right]^{1/2} \quad (58)$$

$$A_0^{(B)} = \left[ \frac{A_0^{(A)}}{c} \right] \pm \left[ \left(\frac{1}{c}\right)^2 u^2(A_0^{(A)}) + \left(\frac{A_0^{(A)}}{c^2}\right)^2 u^2(c) \right]^{1/2} \quad (59)$$

in the absence of further parameter correlation information. Further practical implementation details on the application of the ratio method using the earlier working approximations is provided in the text of Lewis and

Peggs [21] who cite three main reasons for nonlinearity in the pressure/area relationship viz. (i) interference fits effects when high pressure pistons are operated at the lower pressure ranges, (ii) rotational directional influences due to helical scratching on the surfaces of pistons over time, and (iii) errors in the evaluation of the loading force terms, which may be addressed through the introduction of an effective mass term or equivalent model compensation parameter to minimize the errors as already discussed.

Whilst rotational directional influences may be minimized through the reversal of piston spin direction and clockwise and counter-clockwise cross-float measurements that when combined average out any discrepancies one of the main concerns is the nonlinear behavior that is sometimes exhibited when high pressure pistons are utilized at lower operating ranges.

As discussed by Blagojevic and Bajsic [2] this effect has been experimentally observed in pressure balances specifically designed to operate at low pressures and the recommended solution is to use a modified nonlinear area equation for specific pressure balances. This approach is however not suitable for high pressure piston-cylinder pressure balances that are specifically designed, with elastic distortions due to exerted fluid pressure, to exhibit small interference fits between the piston and cylinder at high pressures.

As a result the linear area relationship only strictly holds over a designed limited pressure range and so any deviations from a linear behavior are only present at the extreme of the particular instrument's operating range. Deviations from a linear area behavior where they exist should be examined from the ideal extrapolation to zero pressure, which is not physically realizable in any real instrument, and assessed based on the extent and significance on which it impacts on the effective linear operating range.

## 6 Illustration of method for an oil pressure balance

In order to demonstrate the technique to determine the unknown uncertainties for  $A_0$  and  $\lambda$  of a transfer standard pressure balance when cross-floated against a laboratory standard with known characteristics we apply the direct pressure/area method using the illustrative data in Tables 1–4 for a sequence of 10 pressures from 50 MPa to 500 MPa at 50 MPa steps using the following piston-cylinder data for the LS and TS with the following input data:

Laboratory standard:

$$A_0^{(A)} = 1.96151 \text{ mm}^2 \pm (9.89581 \times 10^{-5}) \text{ mm}^2$$

$$\lambda^{(A)} = (7.25 \pm 0.5) \times 10^{-13} \text{ Pa}^{-1}$$

$$\alpha = (9.10 \pm 0.45) \times 10^{-6} \text{ }^\circ\text{C}^{-1}$$

$$V_s = 0 \text{ m}^3.$$

**Table 1.** Laboratory standard and transfer standard mass piece values where mass piece #1 and mass piece #2 correspond to the piston and bell masses respectively.

Mass #	$m^{(LS)}$ (kg)	$u(m^{(LS)})$ ( $\mu\text{g}$ )	$\rho^{(LS)}$ ( $\text{kg m}^{-3}$ )	$u(\rho^{(LS)})$ ( $\text{kg m}^{-3}$ )	$m^{(TS)}$ (kg)	$u(m^{(TS)})$ ( $\mu\text{g}$ )	$\rho^{(TS)}$ ( $\text{kg m}^{-3}$ )	$u(\rho^{(TS)})$ ( $\text{kg m}^{-3}$ )
1	0.002735	0.000005	8000	20	0.200007	0.000003	7920	20
2	1.497245	0.000005	8000	20	0.800047	0.000001	7914	23.742
3	0	0.000002	8000	25	5.000008	0.000002	7920	25
4	5.00008	0.000002	8000	25	4.999995	0.000002	7920	25
5	5.00004	0.000002	8000	25	4.999985	0.000002	7920	25
6	5.00005	0.000002	8000	25	4.999983	0.000002	7920	25
7	5.00004	0.000002	8000	25	4.999991	0.000002	7920	25
8	5.00005	0.000002	8000	25	4.999979	0.000002	7920	25
9	5.00004	0.000002	8000	25	4.999981	0.000002	7920	25
10	5.00006	0.000002	8000	25	4.999979	0.000002	7920	25
11	5.00004	0.000002	8000	25	4.99997	0.000002	7920	25
12	5	0.000002	8000	25	4.999979	0.000002	7920	25
13	5	0.000002	8000	25	4.999976	0.000002	7920	25
14	5	0.000002	8000	25	4.999974	0.000002	7920	25
15	5	0.000002	8000	25	4.999979	0.000002	7920	25
16	5	0.000002	8000	25	4.999988	0.000002	7920	25
17	5	0.000002	8000	25	4.99999	0.000002	7920	25
18	5	0.000002	8000	25	5.000003	0.000002	7920	25
19	5	0.000002	8000	25	4.999982	0.000002	7920	25
20	5	0.000002	8000	25	4.999996	0.000002	7920	25
21	5	0.000002	8000	25	4.000001	0.000002	7920	25
22	5	0.000002	8000	25	2.00001	0.000002	7920	25
23	3.00005	0.000002	8000	25	2	0.000002	7920	25
24	2.000002	0.000002	8000	25	1.000001	0.000002	7920	25
25	1.000003	0.000002	8000	25	0.5	0.000002	7920	25
26	0.500001	0.000002	8000	25	0.199998	0.000002	7920	25
27	0.3	0.000002	8000	25	0.199999	0.000002	7920	25
28	0.200002	0.000002	8000	25	0.099999	0.000002	7920	25

**Table 2.** Ambient and piston-cylinder temperature results.

Point	$t^{(LS)}$ ( $^{\circ}\text{C}$ )	$p_a$ (kPa)	$h_a$ (%RH/100)	$t_a$ ( $^{\circ}\text{C}$ )	$t^{(TS)}$ ( $^{\circ}\text{C}$ )
1	19.9	86.707	0.63	19	19.33
2	20.0	86.690	0.63	19	19.57
3	20.0	86.670	0.64	19	19.57
4	20.1	86.637	0.66	19	19.82
5	20.1	86.622	0.66	19	19.98
6	20.15	86.606	0.66	19	20.01
7	20.3	86.825	0.67	19	20.33
8	20.4	86.489	0.68	19	20.28
9	20.6	86.442	0.67	19	20.26
10	20.6	86.435	0.67	19	20.28

Transfer standard:

$$\alpha = (1.45 \pm 0.5) \times 10^{-5} \text{C}^{-1}$$

$$C = 4.962871 \text{ mm}$$

$$V_s = 0 \text{ m}^3.$$

Common data:

$$\sigma = (31.2 \pm 0.001) \text{ mN m}^{-1}$$

$$g = (9.7860994 \pm 10^{-7}) \text{ m s}^{-2}$$

$$H = (173.1 \pm 0.2) \text{ mm}$$

with a pressure balance temperature uncertainty of  $\pm 0.015^{\circ}\text{C}$  and by assuming the ambient temperature, pressure and relative humidity can be measured at accuracies of  $\pm 0.5^{\circ}\text{C}$ ,  $\pm 15 \text{ Pa}$  and  $\pm 5 \text{ \%RH}$  respectively.

In specifying the submerged volume  $V_s$  we set it to zero and through the assignment of an arbitrarily small uncertainty such  $\pm 10^{-10} \text{ m}^3$  one may continue to utilize the general measurand model in the asymptotic limit as  $u(V_s) \rightarrow 0$  in the specific case that  $V_s = 0 \text{ m}^3$  with minimal rework of underlying computer code.

The results for the transfer standard pressure balance are shown in Figure 1 which is a plot of the applied pressure  $P$  in MPa against the area in  $\text{mm}^2$  using a computer program implementation for the direct method written in GNU Octave [23]. Further intermediate calculation results are illustrated in Figure 2 and from the plots of the processed measurement results it is clearly seen that a straight line adequately represents the area pressure dependence according to the assumed form of area measurand model  $A = A_0(1 + \lambda P)$ , with the exception of a small nonlinear portion at the extreme of the operating range i.e. for pressures at the lower end of the instruments pressure range.

In general larger cross-floated area uncertainties occur at lower pressures for high pressure piston-cylinder pressure balances and so the use of a linear relationship

**Table 3.** Boolean values of utilized laboratory standard mass pieces used for generated pressures where mass piece #1 and mass piece #2 correspond to the piston and bell masses respectively, and points 1 through 10 correspond to the cross-floated pressures from 50 MPa to 500 MPa respectively.

Point \ Mass	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1
3	0	0	0	0	0	0	0	0	0	0
4	1	1	1	1	1	1	1	1	1	1
5	0	1	1	1	1	1	1	1	1	1
6	0	1	1	1	1	1	1	1	1	1
7	0	0	1	1	1	1	1	1	1	1
8	0	0	1	1	1	1	1	1	1	1
9	0	0	0	1	1	1	1	1	1	1
10	0	0	0	1	1	1	1	1	1	1
11	0	0	0	0	1	1	1	1	1	1
12	0	0	0	0	1	1	1	1	1	1
13	0	0	0	0	0	1	1	1	1	1
14	0	0	0	0	0	1	1	1	1	1
15	0	0	0	0	0	0	1	1	1	1
16	0	0	0	0	0	0	1	1	1	1
17	0	0	0	0	0	0	0	1	1	1
18	0	0	0	0	0	0	0	1	1	1
19	0	0	0	0	0	0	0	0	1	1
20	0	0	0	0	0	0	0	0	1	1
21	0	0	0	0	0	0	0	0	0	1
22	0	0	0	0	0	0	0	0	0	1
23	1	1	1	1	1	1	1	1	1	1
24	0	0	0	0	0	0	0	0	0	0
25	0	0	0	0	0	0	0	0	0	0
26	1	1	1	1	1	1	1	1	1	1
27	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0

**Table 4.** Boolean values of utilized transfer standard mass pieces used for cross-floated pressures where mass piece #1 and mass piece #2 correspond to the piston and bell masses respectively, and points 1 through 10 correspond to the cross-floated pressures from 50 MPa to 500 MPa respectively.

Point \ Mass	1	2	3	4	5	6	7	8	9	10
1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	1	1	1
3	1	1	1	1	1	1	1	1	1	1
4	0	1	1	1	1	1	1	1	1	1
5	0	1	1	1	1	1	1	1	1	1
6	0	0	1	1	1	1	1	1	1	1
7	0	0	1	1	1	1	1	1	1	1
8	0	0	0	1	1	1	1	1	1	1
9	0	0	0	1	1	1	1	1	1	1
10	0	0	0	0	1	1	1	1	1	1
11	0	0	0	0	1	1	1	1	1	1
12	0	0	0	0	0	1	1	1	1	1
13	0	0	0	0	0	1	1	1	1	1
14	0	0	0	0	0	0	1	1	1	1
15	0	0	0	0	0	0	1	1	1	1
16	0	0	0	0	0	0	0	1	1	1
17	0	0	0	0	0	0	0	1	1	1
18	0	0	0	0	0	0	0	0	1	1
19	0	0	0	0	0	0	0	0	1	1
20	0	0	0	0	0	0	0	0	0	1
21	1	1	1	1	1	1	1	1	1	1
22	0	0	0	0	0	0	0	0	0	1
23	0	0	0	0	0	0	0	0	0	1
24	0	0	0	0	0	0	0	0	0	1
25	0	0	0	0	0	0	0	0	0	0
26	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	0	0	0	0	0	0
28	0	0	0	0	0	0	0	0	0	0

extrapolation is seen to have a relatively modest impact on a limited part of the instruments operating range.

A qualitative analysis to study the influence of the zero-pressure area uncertainty  $u(A_0)$  and distortion coefficient uncertainty  $u(\lambda)$  may be performed by considering the fundamental defining equation for the applied pressure  $P = F/A$  from which it follows that the uncertainty in applied pressure  $u(P)$  is

$$\begin{aligned}
 u^2(P) &= \left(\frac{\partial P}{\partial F}\right)^2 u^2(F) + \left(\frac{\partial P}{\partial A}\right)^2 u^2(A) \\
 &= \left(\frac{1}{A}\right)^2 u^2(F) + \left(\frac{-F}{A^2}\right)^2 u^2(A). \tag{60}
 \end{aligned}$$

Considering the underlying area model  $A = A_0(1 + \lambda P)$  it then follows that

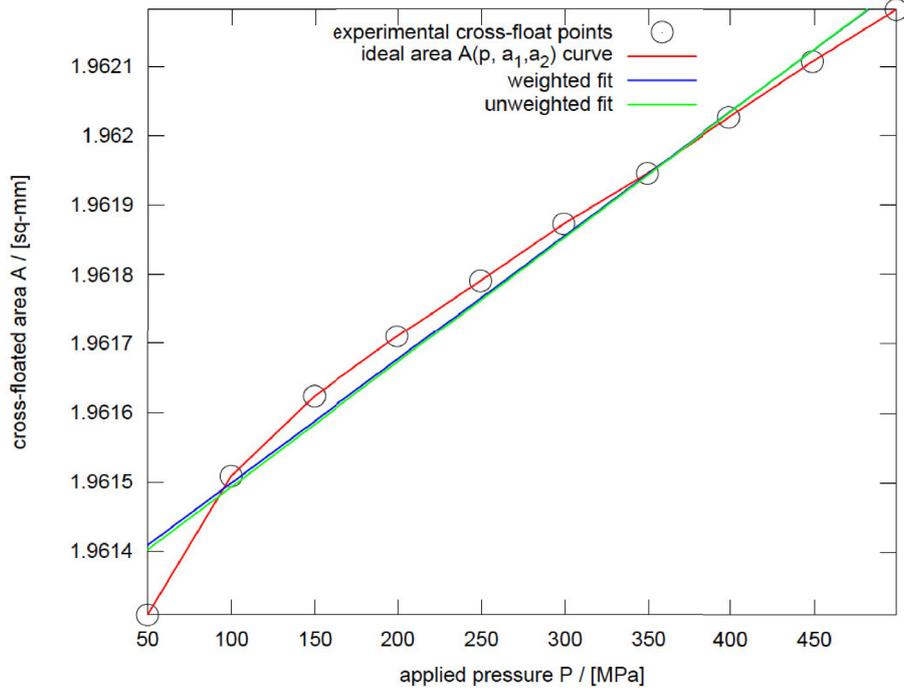
$$u^2(A) = (1 + \lambda P)^2 u^2(A_0) + (A_0 P)^2 u^2(\lambda) \tag{61}$$

which on substituting results in

$$\begin{aligned}
 u(P) &= \left[ \left(\frac{1}{A}\right)^2 u^2(F) + \left(\frac{-F}{A^2}\right)^2 \right. \\
 &\quad \left. \times \left\{ (1 + \lambda P)^2 u^2(A_0) + (A_0 P)^2 u^2(\lambda) \right\} \right]^{1/2}. \tag{62}
 \end{aligned}$$

Restricting the sensitivity analysis to the influence of  $u(A_0)$  and  $u(\lambda)$  only by taking the limit  $u(F) \rightarrow 0$  and noting that  $(1 + \lambda P) = \frac{A}{A_0}$ ,  $A_0 P = \frac{A - A_0}{\lambda}$  and  $\frac{F}{A^2} = \frac{P}{A}$  we have

$$\begin{aligned}
 u(P) &= \left(\frac{F}{A^2}\right) \left\{ (1 + \lambda P)^2 u^2(A_0) + (A_0 P)^2 u^2(\lambda) \right\}^{1/2} \\
 &= \left(\frac{F}{A^2}\right) \left\{ \left(\frac{A}{A_0}\right)^2 u^2(A_0) + \left(\frac{A - A_0}{\lambda}\right)^2 u^2(\lambda) \right\}^{1/2} \tag{63}
 \end{aligned}$$



**Fig. 1.** (Color online) Results of direct pressure/area method calculations for a transfer standard pressure balance calibrated by cross-float to a laboratory standard. The transfer standard pressure balance model parameter is evaluated as  $\mathbf{a} = \{A_0, \lambda\} = \{1.9613197257 \text{ mm}^2, 0.9116463858 \text{ ppm/MPa}\}$  with a corresponding standard uncertainty of  $u(\mathbf{a}) = \{0.0000738752 \text{ mm}^2, 0.1223666979 \text{ ppm/MPa}\}$  i.e.  $u(A_0) = \pm 37.666 \text{ ppm}$  and  $u(\lambda) = \pm 13.422\%$ .

and then by dividing through by the applied pressure  $P$  on both sides yields

$$\begin{aligned} \frac{u(P)}{P} &= \frac{1}{A} \left\{ \left( \frac{A}{A_0} \right)^2 u^2(A_0) + \left( \frac{A - A_0}{\lambda} \right)^2 u^2(\lambda) \right\}^{1/2} \\ &= \left\{ \frac{1}{A^2} \right\}^{1/2} \left\{ \left( \frac{A}{A_0} \right)^2 u^2(A_0) + \left( \frac{A - A_0}{\lambda} \right)^2 u^2(\lambda) \right\}^{1/2} \\ &= \left\{ \left[ \frac{u(A_0)}{A_0} \right]^2 + \left[ \frac{1}{A^2} (A - A_0)^2 \right] \left[ \frac{u(\lambda)}{\lambda} \right]^2 \right\}^{1/2}. \end{aligned} \quad (64)$$

Noting that  $(A - A_0)^2 = A^2 - 2AA_0 + A_0^2$  we have

$$\begin{aligned} \left( \frac{A - A_0}{A} \right)^2 &= \frac{A^2 - 2AA_0 + A_0^2}{A^2} \\ &= 1 - 2\frac{A_0}{A} + \left( \frac{A_0}{A} \right)^2 \\ &= \left( 1 - \frac{A_0}{A} \right)^2. \end{aligned}$$

But since  $A_0 = A/(1 + \lambda P)$  it follows that  $\frac{A_0}{A} = \frac{1}{1 + \lambda P}$  which implies that  $1 - \frac{A_0}{A} = 1 - \frac{1}{1 + \lambda P} = \frac{\lambda P}{1 + \lambda P}$  from which we finally have

$$\left[ \frac{u(P)}{P} \right]^2 = \left[ \frac{u(A_0)}{A_0} \right]^2 + \left( \frac{\lambda P}{1 + \lambda P} \right)^2 \left[ \frac{u(\lambda)}{\lambda} \right]^2. \quad (65)$$

Defining the variables

$$\xi_1 = \frac{u(A_0)}{A_0} \quad (66)$$

$$\xi_2 = \frac{u(\lambda)}{\lambda} \quad (67)$$

$$s = \frac{\lambda P}{1 + \lambda P} \quad (68)$$

$$\eta = \frac{u(P)}{P} \quad (69)$$

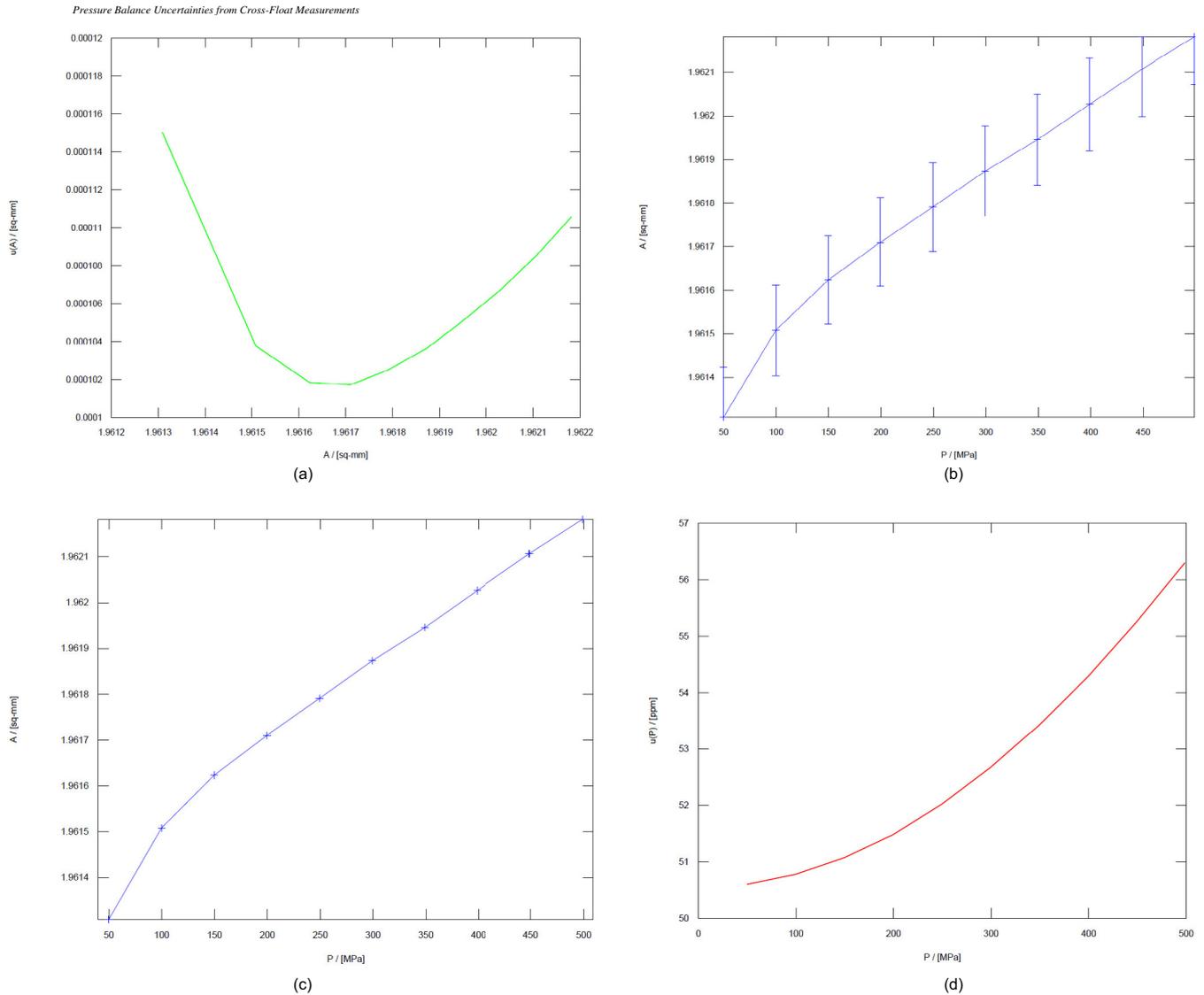
for an equation

$$\eta^2 = \xi_1^2 + s^2 \xi_2^2 \quad (70)$$

then facilitates further sensitivity analysis study in the relative uncertainties of the zero-pressure area and distortion coefficient contribution on the relative uncertainty in applied pressure for the range of operating pressures of the pressure balance as illustrated in Figure 3.

## 7 Extensions to pressure balance models with additional parameters

The analysis for the optimal choice and value of pressure balance measurand parameters may be extended to incorporate additional terms such as the residual mass  $M_e$  term as discussed earlier [1], or in the case where higher order elastic distortion formulations are applicable.

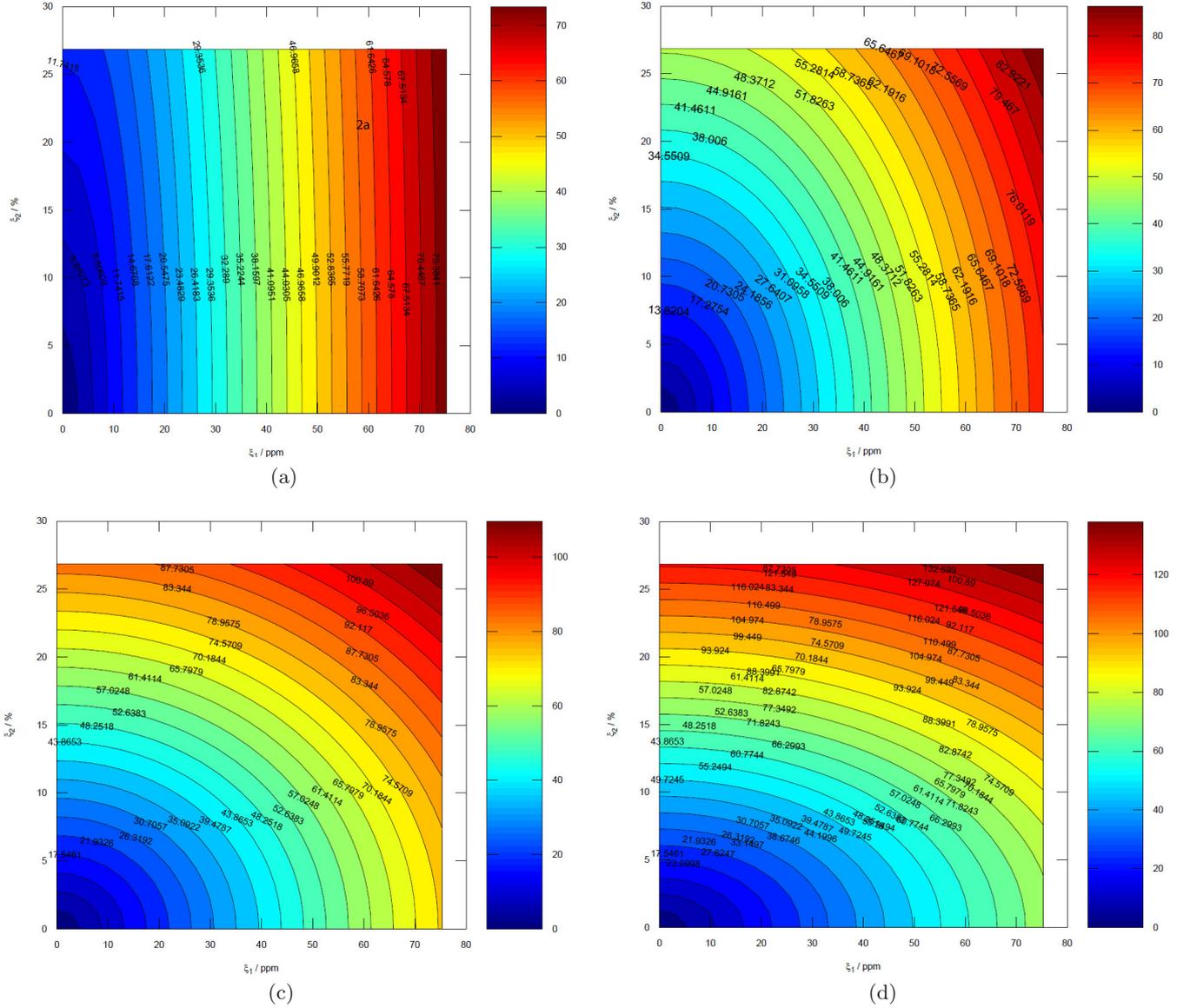


**Fig. 2.** (Color online) Characteristics and associated uncertainty contributors in direct pressure/area method calculations for a transfer standard pressure balance cross-floated against a laboratory standard pressure balance. (a) Uncertainty of transfer standard cross-floated area. (b) Variation of transfer standard uncertainty in cross-floated area with pressure. (c) Actual experimental applied pressure vs. transfer standard cross-floated area. (d) Uncertainty in laboratory standard generated applied pressure.

Examples of such instances include the case of material deformation that exhibits nonlinear deformation under very high pressures for example at operating pressures  $P$  above the piston-cylinder yield strength  $S_y$ . Contemporary pressure balance theory as originally developed by Dadson et al. [3] utilizes linear elasticity theory i.e. where there is a defined linear stress  $\sigma$  strain  $\varepsilon$  relationship of form  $\sigma = E\varepsilon$  where the Young's modulus of elasticity  $E$  is a constant of proportionality from which the traditional linear elastic distortion coefficient formula is derived in terms of geometrical and constituent material property data. Direct engineering analysis outside the elastic limits occurring within the plastic deformation zone where the stress-strain relationship is nonlinear of form  $\sigma = \sigma_0\varepsilon^m$  with material constituent parameters  $\sigma_0$  and  $m$  as discussed in Shigley and Mischke [24] is mathematically

possible in the case of simple idealizations such as perfectly round, straight and concentric piston/cylinders but is more amenable to computational solution using the well known finite element method (FEM) on a practical implementation level [25] either with commercial or in-house research codes.

Other examples that may also be included are where the pressure balance incorporates special features for example in controlled clearance mode i.e. compensatory small control jacket pressures that minimize the elastic deformation and which may necessitate additional model assumptions and parameters [26,27], and in the case where flow models at operating pressures occur that may result in deviations from the commonly used model assumptions such as the validity of Boyles law for ideal gases in the piston-cylinder interface gap [2, 3, 10, 28].



**Fig. 3.** (Color online) Qualitative uncertainty analysis for relative standard uncertainty in zero-pressure area  $\xi_1 = u(A_0)/A_0$  and relative standard uncertainty in distortion coefficient  $\xi_2 = u(\lambda)/\lambda$  contributions to the relative standard uncertainty in applied pressure  $\eta = u(P)/P$  expressed in ppm. The range of  $\xi_i$  ( $i = 1, 2$ ) is chosen as  $0 \leq \xi_i \leq 2u(a_i)/a_i$  where  $\mathbf{a} = \{A_0, \lambda\}$  to encompass the full possible spectrum of potential range of values within the expanded uncertainty using an approximate coverage of  $k = 2$  for a confidence interval at 95% significance. The expected calculated uncertainty values are  $u(A_0)/A_0 = 37.7$  ppm and  $u(\lambda)/\lambda = 13.4\%$  for the illustrative example and it is seen that  $\lim_{\xi_1 \rightarrow 0, \xi_2 \rightarrow 0} \eta(\xi_1, \xi_2) = 0$  correctly converges as expected with  $0 \leq \eta \leq 120$  ppm over the range of applied pressures  $50 \text{ MPa} \leq P \leq 500 \text{ MPa}$ . (a) Results for 50 MPa. (b) Results for 200 MPa. (c) Results for 350 MPa. (d) Results for 500 MPa.

Regardless of the underlying physical or engineering reasons that result in additional model parameters or non-linear effects, in general a least squares regression analysis is necessary in order to compute both the model parameters as well as parameter uncertainties. In order to illustrate the statistical modelling approach for the inclusion of additional parameters we will consider the case of both a higher order elastic distortion coefficient formulation in the form of a quadratic equation viz. an area measurand

model  $A = A_0(1 + \lambda_1 p + \lambda_2 p^2)$  as well as the presence of a residual mass term  $M_e$  using the  $P$ -method as our implementation approach.

For the case of a laboratory standard (LS) used to calibrate a transfer standard (TS) it follows that the reference value for the pressure is that of the LS to which the inferred parameters of the TS must be optimized to minimize an error term.

Following the approach of Press et al. [8] a merit function  $\chi^2(y|\mathbf{a})$  must be constructed, where according to our modelling approach  $\mathbf{a} = [A_0, \lambda_1, \lambda_2, M_e]^T$  is a vector of the model parameter components, which in our case will take the form

$$\chi^2(P|\mathbf{a}) = \sum_{i=1}^q \left[ \frac{1}{A_0(1+\lambda_1 p + \lambda_2 p^2)} \left\{ \sum_{j=1}^n m_j \left(1 - \frac{\rho_a}{\rho_j}\right) g + M_e g + \sigma C - V_s (\rho_f - \rho_a) g \right\} - P_i \right]^2 \quad (71)$$

noting that the applied pressure  $P$  is simply the difference  $p - p_a$  between the absolute pressure  $p$  and the atmospheric pressure  $p_a$ .

Since this particular merit function is not constructed as a linear combination of underlying, possibly nonlinear in the argument  $x$ , basis functions  $X_i$  i.e. of the form  $\chi^2 = \sum_{i=1}^k a_k X_k(x)$  as occurred in the original GLS of Sutton et al. [1] i.e. linear combinations in the parameters  $\beta_1 = M_e$ ,  $\beta_2 = A_0$  and  $\beta_3 = \lambda A_0$  it is considered nonlinear as per standard statistical modelling methodology, and for which a linearization of form

$$\chi^2(\mathbf{a}) \approx \gamma - \mathbf{a} \cdot \mathbf{a} + \frac{1}{2} \mathbf{a} \cdot \mathbf{D} \cdot \mathbf{a} \quad (72)$$

is necessary in order to perform the regression analysis. In order to calculate the matrix  $\mathbf{D}$  we observe that the general form of the merit function is  $\chi^2 = \sum_{i=1}^q \left[ \frac{y_i - y(x_i|\mathbf{a})}{\sigma_i} \right]^2$  where the weighting factor  $\sigma_i$  for the  $i$ th data point is  $\sigma_i = \text{Var}(y_i - y(x_i|\mathbf{a}))$ . The derivation for the optimal choice of measurand parameter  $\mathbf{a}$  follows from the observation that at the optimum  $\mathbf{a}$  that the partial derivatives  $\partial \chi^2 / \partial a_j$ ,  $j = 1, \dots, n$  where  $n$  is the number of components of  $\mathbf{a}$  equal zero i.e. a local minima occurs.

At this point it may seem advantageous to apply the GUM Supplement 1 or GUM-S1 [29] in order to determine the variances in the components  $a_i$  of the parameter  $\mathbf{a}$  by a direct Monte Carlo simulation using sampled values from the measurement data. The general idea in attempting to apply the GUM-S1 method for regression analysis parameter uncertainty estimation is to utilize the actual experimental mass pieces, temperature, air density values etc. that were recorded in a cross-float measurement of the laboratory standard against the transfer standard to compute the applicable parameter expected values i.e. the means  $\mu_i$ , and to then repeat the calculations by utilizing sampled values from the population distribution of the measurement data, and thus build up a frequency distribution of the parameter components from which the parameter variances may be computed.

Unfortunately a straightforward application of GUM-S1 is not in general suitable as discussed and investigated in a recent paper by Elster and Toman [30] where the principle objection encountered in utilizing this technique as a general approach is that a regression problem can not be uniquely transformed into an equivalent measurement

model, and where it was determined that results in both methods are only equivalent in the special case for a linear regression model with known variance and in the absence of prior knowledge about the measurand.

As a result it is seen that a GUM-S1 type approach to a regression analysis is not suitable for parameter variance estimation since the underlying merit function is nonlinear, the variance of the measurand is known since it defined by the laboratory standard calculated uncertainty, and existing knowledge of the measurand i.e. the generated pressure is present.

In order to perform the nonlinear regression analysis and parameter variance estimation one may either apply a Levenberg-Marquardt optimization of the merit function as discussed in Press et al. [8] or alternately on a more powerful generalized basis perform a Markov Chain Monte Carlo (MCMC) simulation as recommended by Elster and Toman [30] and discussed in Forbes [31] for a more complete Bayesian uncertainty analysis as an avenue for future exploration.

Once the applicable parameter  $\mathbf{a} = [A_0, \lambda_1, \lambda_2, M_e]^T$  and its components uncertainties are computed by the MCMC simulation one may then utilize either the approximations within the GUM methodology as outlined earlier or alternately the GUM Supplement 1 methodology for more refined uncertainty calculations using extensions to implicit models as discussed in Cox and Harris [17, 32] to incorporate for example physical aspects such as fluid density dependence on operating pressure, surface tension, and fluid heating effects that without simplifying approximations are only amenable by direct Monte Carlo simulation.

## 8 Conclusions

In this paper we have developed the methodology to consistently perform the uncertainty analysis for a piston-cylinder pressure balance, applicable to fluid mediums of both liquids such as oils as well as to gases, characterized by the zero-pressure effective area  $A_0$  and distortion coefficient  $\lambda$ .

Both the zero-pressure area uncertainty as well as the distortion coefficient uncertainty were derived by application of the developed methodology for common existing working approximations, and which represents an improvement on existing disparate and incomplete distortion coefficient uncertainty analyses. The uncertainties in  $\lambda$  and  $A_0$  for a transfer standard were determined when utilizing the method of cross-floating with another reference pressure balance standard whose parameters and associated uncertainties are known for both the direct pressure/area as well as indirect ratio methods.

Extensions of the developed method for a pressure balance's measurand model parameter uncertainty were outlined to encompass more refined accuracy specifications and to incorporate nonlinear area pressure behaviour without any limiting assumptions and approximations for future work.

*Acknowledgements.* This work was performed with funding provided by the Department of Trade and Industry on behalf of the South African government for technical infrastructure agencies.

## Appendix A: Derivation of the covariance of a product of two non-independent random variables

Our objective is to determine the covariance  $\text{Cov}(X, Y)$  between two real-valued random variables  $X$  and  $Y$  which is by definition

$$\begin{aligned} \text{Cov}(X, Y) &= E[(X - E(X))(Y - E(Y))] \\ &= E[XY] - E[X]E[Y] \end{aligned} \quad (\text{A.1})$$

in the case where  $X$  and  $Y$  have finite second moments i.e. where the  $n$ th moment of  $X$  about a value  $c$  is defined as  $\mu'_n = \int_{-\infty}^{+\infty} (x-c)^n f(x)dx$ , and where the expectation is defined as  $E[X] = \int_{-\infty}^{+\infty} xf(x)dx$  for a probability density function  $f(x)$  as per standard statistical definitions.

In the derivation which follows we will suppose that that the random variable  $X$  or  $X_1$  represents the zero-pressure effective area  $A_0$  and the random variable  $Y$  or  $X_2$  represents the distortion coefficient  $\lambda$ . The physical significance of the specification of a finite second moment is that it equals the variance  $\text{Var}(X) = E[(X - \mu)^2]$  and the boundedness thus allows for finite physical uncertainties in  $A_0$  and  $\lambda$ .

We follow the approach of Wilkinson [5] for the derivation of the variance of  $k$  products of random variables in the most general case in order to determine the covariance between two random variables. Utilizing the notation

$$E(x_{ji}) = X_{ji} \quad (\text{A.2})$$

$$G(x_{ji}) = \frac{V(x_{ji})}{X_{ji}^2} \quad (\text{A.3})$$

$$\Delta x_{ji} = x_{ji} - X_{ji} \quad (\text{A.4})$$

$$\delta x_{ji} = \frac{\Delta x_{ji}}{X_{ji}} \quad (\text{A.5})$$

$$D(x_{1i}, x_{2i}) = \frac{\text{Cov}(x_{1i}, x_{2i})}{X_{1i}X_{2i}} \quad (\text{A.6})$$

with the shorthand of  $V(\cdot) = \text{Var}(\cdot)$  for the variance for two sets  $(x_{11}, x_{21})$  and  $(x_{12}, x_{22})$  of two independent random variables where  $E$  is the expectation as discussed in Wilkinson [5] and where

$$D(z_1, z_2, \dots, z_k) = E[(\delta z_1)(\delta z_2) \cdots (\delta z_k)]. \quad (\text{A.7})$$

In the special case for two random variables i.e.  $k = 2$  Wilkinson [5] was able to derive the exact and approxi-

mate formulae for the variance of the product  $x_1x_2$  as

$$\begin{aligned} V(x_1x_2) &= X_1^2X_2^2 \left[ G(x_1) + G(x_2) + 2D(x_1, x_2) \right. \\ &\quad \left. + 2D(x_1, x_1, x_2) + 2D(x_1, x_2, x_2) \right. \\ &\quad \left. + D(x_1, x_1, x_2, x_2) - \{D(x_1, x_2)\}^2 \right] \end{aligned} \quad (\text{A.8})$$

$$\approx X_1^2X_2^2 [G(x_1) + G(x_2) + 2D(x_1, x_2)]. \quad (\text{A.9})$$

As discussed in [5] all terms for the calculation of  $D$  with more than two arguments involve higher order moments, and as a result the formula for the variance  $V(x_1x_2)$  is only feasible when the approximation is utilized. Expanding out explicitly in the random variables  $x$  and  $y$  for the product  $w = xy$  we have

$$\begin{aligned} V(xy) &\approx X^2Y^2 [G(x) + G(y) + 2D(xy)] \\ &= X^2Y^2 \left[ \frac{V(x)}{X^2} + \frac{V(y)}{Y^2} + 2D(x, y) \right] \\ &\Leftrightarrow 2 \frac{\text{Cov}(x, y)}{XY} = \frac{V(xy)}{X^2Y^2} - \frac{V(x)}{X^2} - \frac{V(y)}{Y^2} \end{aligned} \quad (\text{A.10})$$

by applying the definition  $D(xy) = \text{Cov}(x, y)/(XY)$ , from which we finally obtain

$$\begin{aligned} \text{Cov}(x, y) &= \frac{1}{2} \frac{1}{XY} [V(xy) - Y^2V(x) - X^2V(y)] \\ &= \frac{1}{2} \left[ \frac{V(xy)}{XY} - \frac{y}{x}V(x) - \frac{x}{y}V(y) \right]. \end{aligned} \quad (\text{A.11})$$

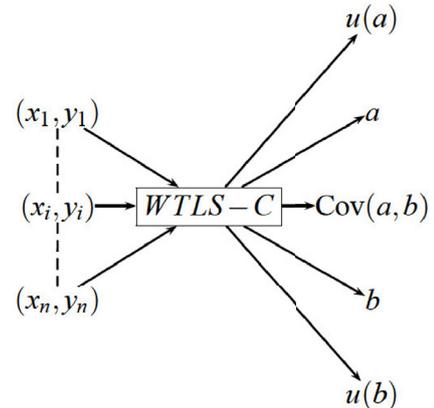
In order to assess the accuracy of the approximation one may consider the well known Cauchy-Schwarz inequality

$$|\text{Cov}(X, Y)| \leq \sqrt{V(X)V(Y)}. \quad (\text{A.12})$$

In the special case for the product of independent random variables Wilkinson [5] derived the formula

$$V(xy) = X^2V(y) + Y^2V(x) + V(x)V(y). \quad (\text{A.13})$$

The covariances of the product of two variables  $xy$  are necessary in order to apply the GUM formula for the uncertainty of the straight line fit parameters from the WTLS-C program as illustrated schematically below:



### Appendix B: Calculation of full explicit higher order gum uncertainty expression

The GUM [6] specifies the recommended calculation for the uncertainty of a measurand defined by an equation

$$y = f(x_1, \dots, x_n) \tag{B.1}$$

where  $n$  is the number of independant random variables and where the full expression for the uncertainty in  $y$  is defined as per equation (7), which contains covariances as well as higher order nonlinear terms. Frequently this calculation is approximated utilizing just the linear and first order approximations as

$$u^2(f) \approx \sum_{i=1}^n \left( \frac{\partial f}{\partial x_i} \right)^2 u^2(x_i) + 2 \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{\partial f}{\partial x_i} \frac{\partial f}{\partial x_j} u(x_i, x_j). \tag{B.2}$$

Utilizing a computer algebra system such as Maxima [22] can result in a full explicit expression with the computer code as per Listing 1 with results for a measurand composed of the the product of two random variables i.e. by setting  $x = x_1$  and  $y = x_2$  we have

$$f = x_1 x_2 \tag{B.3}$$

$$\therefore u^2(f) = V(x_1) x_2^2 + 2x_1 \text{Cov}(x_1, x_2) x_2 + x_1^2 V(x_2) + V(x_1) V(x_2) \tag{B.4}$$

$$\Rightarrow \text{Cov}(x_1, x_2) = \frac{1}{2} \left[ \frac{V(x_1 x_2)}{x_1 x_2} - \frac{x_2}{x_1} V(x_1) - \frac{x_1}{x_2} V(x_2) - \frac{V(x_1) V(x_2)}{x_1 x_2} \right]. \tag{B.5}$$

It is observed that this GUM derived expression differs from that of [5] which calculates the covariance as

$$\text{Cov}(x, y) = \frac{1}{2} \left[ \frac{\text{Var}(xy)}{XY} - \frac{y}{x} \text{Var}(x) - \frac{x}{y} \text{Var}(y) \right]. \tag{B.6}$$

We remark that the expression developed by [5] is developed from statistical first principles without any limiting approximating assumptions such as occurs in the GUM which is strictly valid for a local linearization of the measurand e.g.

$$f(x, y) \approx f(a, b) + \left. \frac{\partial f(x, y)}{\partial x} \right|_{x=a, y=b} (x - a) + \left. \frac{\partial f(x, y)}{\partial y} \right|_{x=a, y=b} (y - b) \tag{B.7}$$

or equivalently [33] the assertion that the expectation of the measurand function equals the measurand function of the expectation of the measurand i.e.

$$E[f(\mathbf{x})] = f(E[\mathbf{x}]) \tag{B.8}$$

and as a result we opt for the use of the covariance formula of Wilkinson [5] instead of attempting to apply the GUM [6] for a purpose it was not originally developed for.

### Appendix C: Review of algorithm implementation for weighted least-squares straight line fit

This section presents a condensed summary for convenience of the least-squares algorithm published in Krystek and Anton [19] for the fitting of a straight line of the form  $y = ax + b$  to  $\mathbf{x}$  and  $\mathbf{y}$  data which in our case corresponds to applied pressure and area data.

In the paper of [19] the underlying data is fitted with parameters  $\alpha$  and  $p$  for convenience where  $\alpha$  is the angle of the straight line from the horizontal  $x$ -axis measured in an anti-clockwise direction and  $p$  is the perpendicular distance from the straight line to the origin of the  $xy$  coordinate system.

For  $n$  data points inputs with underlying data of  $x_k$  and  $y_k$  and associated variances  $\text{Var}(x_k)$ ,  $\text{Var}(y_k)$  and covariances  $\text{Cov}(x_k, y_k)$  apply the following steps:

1. Set

$$u_{x,k} = \sqrt{\text{Var}(x_k)} \tag{C.1}$$

$$u_{y,k} = \sqrt{\text{Var}(y_k)} \tag{C.2}$$

$$\rho_k = \frac{\text{Cov}(x_k, y_k)}{\sqrt{\text{Var}(x_k) \text{Var}(y_k)}}. \tag{C.3}$$

2. Set

$$v_k = y_k \cos \alpha - x_k \sin \alpha - p \tag{C.4}$$

$$g_k = u_{x,k}^2 \sin^2 \alpha - 2\rho_k u_{x,k} u_{y,k} \sin \alpha \cos \alpha + u_{y,k}^2 \cos^2 \alpha. \tag{C.5}$$

3. Set

$$u = \left( \frac{1}{n} \sum_{k=1}^n \frac{1}{g_k} \right)^{-1/2} \tag{C.6}$$

$$w_k = \frac{u^2}{g_k}. \tag{C.7}$$

4. Calculate

$$\bar{x} = \frac{1}{n} \sum_{k=1}^n w_k x_k \tag{C.8}$$

$$\bar{y} = \frac{1}{n} \sum_{k=1}^n w_k y_k. \tag{C.9}$$

5. Set

$$p = \bar{y} \cos \alpha - \bar{x} \sin \alpha. \tag{C.10}$$

6. Form the merit function

$$\chi^2 = \sum_{k=1}^n \frac{v_k^2}{g_k}. \tag{C.11}$$

7. Optimize the merit function in the single variable  $\alpha$  i.e. solve the nonlinear equation [34]

$$\frac{\partial \chi^2}{\partial \alpha} = 0 \quad (\text{C.12})$$

for a solution  $\alpha_0$  using either Newton's method to solve an equation  $f(x) = 0$  with an initial approximation  $p_0$  that generate a sequence of solutions  $\{p_n\}_{n=0}^{\infty}$  with

$$p_n = p_{n-1} - \frac{f(p_{n-1})}{f'(p_{n-1})} \quad (\text{C.13})$$

if derivative data  $\left. \frac{df}{dx} \right|_{x=p_{n-1}} = f'(p_{n-1})$  is available or alternately through the Secant method with a sequence of solutions generated with

$$p_n = p_{n-1} - \frac{f(p_{n-1})(p_{n-1} - p_{n-2})}{f(p_{n-1}) - f(p_{n-2})} \quad (\text{C.14})$$

if analytical derivative data is not available and a pure numerical approach is necessary.

8. Calculate

$$p_0 = \bar{y} \cos \alpha_0 - \bar{x} \sin \alpha_0. \quad (\text{C.15})$$

9. Calculate if  $\alpha \neq (2m+1)\frac{\pi}{2}$ ,  $m \in \mathbf{Z}_0^+$  i.e. in the case that the line is not vertical the straight line  $y = ax + b$  parameters as

$$a = \tan \alpha_0 \quad (\text{C.16})$$

$$b = \frac{p_0}{\cos \alpha_0}. \quad (\text{C.17})$$

10. Set with  $\alpha = \alpha_0$

$$v_{\alpha,k} = -(y_k \cos \alpha + x \cos \alpha) \quad (\text{C.18})$$

$$v_{\alpha\alpha,k} = -y_k \cos \alpha + x \sin \alpha \quad (\text{C.19})$$

and

$$g_k = \frac{u_{x,k}^2 + u_{y,k}^2}{2} - \frac{u_{x,k}^2 - u_{y,k}^2}{2} \cos(2\alpha) - \rho_k u_{x,k} u_{y,k} \sin(2\alpha) \quad (\text{C.20})$$

$$g_{\alpha,k} = (u_{x,k}^2 - u_{y,k}^2) \sin(2\alpha) - 2\rho_k u_{x,k} u_{y,k} \cos(2\alpha) \quad (\text{C.21})$$

$$g_{\alpha\alpha,k} = 2(u_{x,k}^2 + u_{y,k}^2) - 4g_k. \quad (\text{C.22})$$

11. Form the terms

$$\chi_{\alpha\alpha}^2 = \sum_{k=1}^n \left[ \frac{2(v_{\alpha,k}^2 + v_k v_{\alpha\alpha,k})}{g_k} - \frac{4v_k v_{\alpha,k} g_{\alpha,k} + v_k^2 g_{\alpha\alpha,k}}{g_k^2} + \frac{2v_k^2 g_{\alpha,k}^2}{g_k^3} \right] \quad (\text{C.23})$$

$$\chi_{p\alpha}^2 = -2 \sum_{k=1}^n \left[ \frac{v_{\alpha,k}}{g_k} - \frac{v_k g_{\alpha,k}}{g_k^2} \right] \quad (\text{C.24})$$

$$\chi_{pp}^2 = \frac{2n}{u^2}. \quad (\text{C.25})$$

12. Construct the matrix

$$\mathbf{C} = \frac{2}{\chi_{pp}^2 \chi_{\alpha\alpha}^2 - (\chi_{p\alpha}^2)^2} \begin{bmatrix} \chi_{\alpha\alpha}^2 & -\chi_{p\alpha}^2 \\ -\chi_{p\alpha}^2 & \chi_{pp}^2 \end{bmatrix}. \quad (\text{C.26})$$

13. Determine the corresponding variance and covariance parameter terms from

$$\mathbf{C} = \begin{bmatrix} \text{Var}(p) & \text{Cov}(p, \alpha) \\ \text{Cov}(p, \alpha) & \text{Var}(\alpha) \end{bmatrix}. \quad (\text{C.27})$$

14. Calculate the parameter uncertainties as

$$\text{Var}(a) = \frac{\text{Var}(\alpha)}{\cos^4 \alpha} \quad (\text{C.28})$$

$$\text{Var}(b) = \frac{1}{\cos^4 \alpha} \left[ \text{Var}(\alpha) p^2 \sin^2 \alpha + \text{Var}(p) \cos^2 \alpha + \text{Cov}(\alpha, p) p \sin(2\alpha) \right] \quad (\text{C.29})$$

$$\text{Cov}(a, b) = \frac{1}{\cos^4 \alpha} [\text{Var}(\alpha) p \sin \alpha + \text{Cov}(\alpha, p) \cos \alpha]. \quad (\text{C.30})$$

The reader should consult the paper of Krystek and Anton [19] for further background details in the case of the presence of correlation for the weighted least squares curve fit i.e. the WTLS-C algorithm and note that in the absence of correlation i.e.  $\rho_k = 0 \forall k \in [1, \dots, n]$  that the WTLS-C correctly reduces to that of the 'weighted total least squares' i.e. WTLS algorithm published in an earlier paper by the same authors [35].

The above outline provides a means to implement the WTLS-C algorithm, either in MATLAB for which a publicly available routine is available for download [36], or in another computer language such as Fortran, C/C++ or Python where the main requirement to implement the algorithm is a minimization routine.

Due to the fact the method reduces the two parameter merit function  $\chi^2(p, \alpha)$  optimization to a one dimensional minimization search one therefore simply requires a one dimensional minimization search i.e.  $\min_{\alpha} \chi^2(\alpha)$ , for which a golden section numerical routine implementation utilizing only functional evaluations is accessible without recourse to more specialist optimization techniques.

```

/* Copyright (c) Vishal Ramnath */
/* National Metrology Institute of South Africa */
kill(all); /* initialize workspace */
n: read("Enter number of variables and press Enter: ");
"Measurand input example"$
"    f: x[2] + x[1]*x[3]"$
"Then press Shift + Enter"$
f: read("Enter f: ");
cov: genmatrix(C, n, n); /* build up covariance matrix */
                        /* variance in symbolic form */
v: sum(diff(f,x[i],1)^2*u[i]^2, i, 1, n)
+ sum(
    sum( ((1/2)*(diff(f, x[i], 1, x[j], 1))^2
+ diff(f, x[i], 1)
    *diff(f, x[i], 1, x[j], 2))
    *u[i]^2*u[j]^2,
    j, 1, n),
    i, 1, n
)
+ 2*sum(
    sum(diff(f, x[i], 1)
    *diff(f, x[j], 1)
    *cov[i,j],
    j, i + 1, n),
    i, 1, n - 1
);

perform_calculations: read("yes press 1 or press 0 to exit: ");

if ( is(equal(perform_calculations, 1)) ) then
(
    uy: sqrt(v),
    for i: 1 step 1 thru n do
    (
        xval[i]: read("x[" , i, "] = "),
        uval[i]: read("u[" , i, "] = ")
    ),
    Cval: entermatrix(n, n),
    vval: ev(v, x = xval, u = uval, C = Cval),
    uyval: sqrt(vval)
);

```

Listing 1. Maxima/Lisp Implementation of GUM Formula for an Explicit Univariate Measurand

## References

1. C.M. Sutton, M.P. Fitzgerald, W. Giardini, A Method of analysis for cross-floats between pressure balances, *Metrologia* **42**, S212–S215 (2005)
2. B. Blagojevic, I. Bajsic, Determination of the Effective Area of a Gas-Pressure Balance for Low Pressures, in *XVII IMEKO World Congress*, edited by K. Havrilla (2003)
3. R.S. Dadson, S.L. Lewis, G.N. Peggs, *The Pressure Balance: Theory and Practice* (HMSO, London, 1982)
4. A.B. Forbes, P.M. Harris, Estimation algorithms in the calculation of effective area of pressure balances, *Metrologia* **36**, 689–692 (1999)
5. J.W. Wilkinson, *Exact Expressions for the Covariances Between Products of Randoms Variables*, Technical report ORNL-TM-1977, Oak Ridge National Laboratory, Atomic Energy Commission, USA, 1967
6. BIPM, IEC, IFCC, ISO, IUPAC, OIML, Guide to the Expression of Uncertainty in Measurement, 2nd edn., BIPM (1995), ISBN 92-67-10188-9
7. B. Efron, G. Gong, A Leisurely look at the bootstrap, the Jackknife, and cross-validation, *Am. Stat.* **37**, 36–48 (1983)
8. W.H. Press, S.A. Teukolsky, W.T. Vetterling, B.P. Flannery, *Numerical Recipes: The Art of Scientific Computing*, 3rd edn. (Cambridge University Press, 2007)
9. M.G. Cox, B.R.L. Siebert, The use of a Monte Carlo method for evaluating uncertainty and expanded uncertainty, *Metrologia* **43**, S178–S188 (2006)
10. V. Ramnath, Comparison of the GUM and Monte Carlo measurement uncertainty techniques with application to effective area determination in pressure standards, *Int. J. Metrol. Qual. Eng.* **1**, 51–57 (2010)
11. D. Zvizdic, L.G. Bermanec, G. Bonnier, E. Renaot, Uncertainty Budget of Pressure Balance Effective Area Determined by Comparison Method, in *XVII IMEKO World Congress* (2003)
12. A. Picard, R.S. Davis, M. Glaser, K. Fujii, Revised formula for the density of moist air (CIPM-2007), *Metrologia* **45**, 149–155 (2007)
13. E.C. Morris, K.M.K. Fen, *The Calibration of Weights and Balances*, Monograph 4: NML Technology Transfer Series, 3rd edn. (CSIRO National Measurement Laboratory, 2003)
14. D. Tabor, *Gases, Liquids and Solids*, 2nd edn. (Cambridge University Press, 1985)
15. P.J. Mohr, B.N. Taylor, D.B. Newell, CODATA recommended values of the fundamental physical constants: 2006, *Rev. Mod. Phys.* **80** (2008)
16. I. Kocas, *Technical Protocol of EURAMET.M.P-K13 (500 MPa)*, Technical report, TUBITAK UME, 2009
17. M.G. Cox, P.M. Harris, *Software Support for Metrology Best Practice Guide No. 6 – Uncertainty Evaluation*, Technical report, NPL, UK, 2004
18. B.R.L. Siebert, P. Ciarlini, D. Sibold, Monte Carlo Study on Logical and Statistical Correlation, in *VII IMEKO World Congress*, edited by P. Ciarlini, E. Filipe, A.B. Forbes, F. Pavese, C. Perruchet, B.R.L. Siebert (2005)
19. M. Krystek, M. Anton, A least-squares algorithm for fitting data points with mutually correlated coordinates to a straight line, *Meas. Sci. Technol.* **22**, 1–9 (2011)
20. P. Penfield, Principle of Maximum Entropy, in *Information, Entropy and Computation* (2010), [www.mit.edu](http://www.mit.edu), <http://www.mtl.mit.edu/Courses/6.050/2010/notes/chapter9.pdf>
21. S. Lewis, G. Peggs, *The Pressure Balance: A Practical Guide to its Use*, 2nd edn. (HMSO, London, 1992), ISBN 0114800618
22. W. Schelter, *Maxima – A Computer Algebra System (Version 5.23.2)* (Free Software Foundation, 2011), <http://maxima.sourceforge.net/>
23. J.W. Eaton, *GNU Octave Manual*, Network Theory Limited (2002), ISBN 0-9541617-2-6, [www.gnu.org/software/octave](http://www.gnu.org/software/octave)
24. J.E. Shigley, C.R. Mischke, *Mechanical Engineering Design*, 6th edn. (Metric Edition, McGraw-Hill, 2003)
25. S. Dogra, S. Yadav, A.K. Bandyopadhyay, Computer simulation of a 1.0 GPa piston-cylinder assembly using finite element analysis (FEA), *Measurement* **43**, 1345–1354 (2010)
26. A.K. Bandyopadhyay, D.A. Olson, Characterization of a compact 200 MPa controlled clearance piston gauge as a primary pressure standard using the heydemann and welch method, *Metrologia* **43**, 573–582 (2006)
27. P.L.M. Heydemann, B.E. Welch, Piston Gauges, in *Experimental Thermodynamics, Experimental Thermodynamics of Non-Reacting Fluids*, edited by B. Le Neindre, B. Vodar (Butterworth and Co., 1975), Chap. 4, pp. 147–202
28. F.M. White, *Viscous Fluid Flow*, 2nd edn. (McGraw-Hill, 1991)
29. JCGM/WG1, *Evaluation of Measurement Data – Supplement 1 to the Guide to the Expression of Uncertainty in Measurement – Propagation of Distributions using a Monte Carlo Method*, 1st edn., Joint Committee for Guides in Metrology (2007)
30. C. Elster, B. Toman, Bayesian uncertainty analysis for a regression model versus application of GUM supplement 1 to the least-squares estimate, *Metrologia* **48**, 233–240 (2011)
31. A.B. Forbes, Nonlinear Least Squares and Bayesian Inference, in *Advanced Mathematical and Computational Tools in Metrology and Testing*, edited by F. Pavese, M. Bar, A.B. Forbes, J.M. Linares, C. Perruchet, N.F. Zhang (World Scientific Publishing Company, 2009), pp. 104–112
32. M.G. Cox, P.M. Harris, *Software Specifications for Uncertainty Evaluation*, Technical report, NPL, UK, 2006
33. W. Bich, Uncertainty Evaluation by Means of a Monte Carlo Approach, in *BIPM Workshop 2 on CCRI (II) Activity Uncertainties and Comparisons* (2008), [www.bipm.org](http://www.bipm.org), <http://www.bipm.org/wg/CCRI%28II%29/WORKSHOP%28II%29/Allowed/2/Bich.pdf>
34. R.L. Burden, J.D. Faires, *Numerical Analysis* (Brooks/Cole, 2001), ISBN 0-534-38216-9
35. M. Krystek, M. Anton, A weighted total least-squares algorithm for fitting a straight line, *Meas. Sci. Technol.* **18**, 3438–3442 (2007)
36. M. Krystek, M. Anton, Weighted total least squares for mutually correlated coordinates (2011), [www.mathworks.com](http://www.mathworks.com)