

December 1995

A comparison of methods used for the calculation of effective area in the calibration of pressure balances

A B Forbes and P M Harris
Centre for Information Systems Engineering
National Physical Laboratory
Teddington
Middlesex
United Kingdom
TW11 0LW

ABSTRACT

The calibration of a pressure balance involves taking measurements of pressure, applied load and temperature, and estimating from these the values of parameters that define the effective area of the balance as a function of pressure. In this report we discuss and assess various approaches to analysing the measured data to obtain these estimates. In practice, the so-called P - and ΔP -methods are used for the analysis, and we indicate properties of these methods viewed as linear estimators of the effective area parameters. The stated properties are verified by applying the methods considered to a variety of simulated data sets. Furthermore, the numerical simulations provide a means for quantifying and comparing the performance of the methods.

© Crown copyright 1995
Reproduced by permission of the Controller of HMSO

ISSN 1361-407X

National Physical Laboratory
Teddington, Middlesex, United Kingdom, TW11 0LW

Extracts from this report may be reproduced
provided that the source is acknowledged

Approved on behalf of the Managing Director of NPL by
Mr A J Marks, Director, Centre for Information Systems Engineering

CONTENTS

	Page
1 SUMMARY	
2 INTRODUCTION	4
3 MODELS	5
3 THE PRESSURE BALANCE AS A PRESSURE GENERATOR	5
3.2 CALIBRATION OF A PRESSURE BALANCE	5
4 ESTIMATORS	6
LINEAR LEAST-SQUARES METHOD . . .	6
4.2 P-METHOD	7
ΔP-METHOD	8
MEASUREMENT STRATEGY	8
5 DATA GENERATION	9
5.1 GENERATING DATA TO SATISFY EXACTLY THE PRESSURE EQUATION	9
LEAST-SQUARES DATA GENERATOR	10
5.3 GENERALIZED DISTANCE REGRESSION DATA GENERATOR .	11
5.4 GENERATING TYPICAL DATA	12
5.5 COMPARISON OF RESULTS	13
6 NUMERICAL SIMULATIONS	14
6.1 EXACT DATA	14
6.2 SIMULATING THE P- AND ΔP-METHODS	15
6.3 THE EFFECT OF USING DIFFERENT PRIOR KNOWLEDGE IN THE ΔP-METHOD	16
6.4 DATA WITH RANDOM MEASUREMENT ERROR	16
6.4.1 Results for data with normally distributed random error	16
6.4.2 Results for data for which the LLS estimator returns the true parameter values	17
6.4.3 Results for data for which the GDR estimator returns the true parameter values	17
6.5 DATA WITH RANDOM MEASUREMENT ERROR AND SYSTEMATIC ERROR	18
6.6 DATA BASED ON A UNIFORM DISTRIBUTION OF PRESSURE VALUES	18
7 CONCLUSIONS	19

8 ACKNOWLEDGEMENTS	20
9 REFERENCES	21
A NUMERICAL RESULTS	22
B SOFTWARE LISTINGS . . .	29
B.1 LINEAR LEAST-SQUARES METHOD	29
<i>P</i> -METHOD	31
B.3 ΔP-METHOD	33
B.4 GENERATION OF EXACT DATA	36
GDR DATA GENERATION	37

1 SUMMARY

The essential structure of a pressure balance is a piston-cylinder combination. The upward force due to the pressure in the system is balanced against the downward gravitational force generated by known masses acting on the piston. The *effective area* of the piston-cylinder assembly is the quantity by which the applied force must be divided to derive the applied pressure. It is the cross-sectional area of a cylindrical surface, the neutral surface, at a position approximately halfway between the piston and cylinder. Consequently, the effective area is a function of the dimensions of both the piston and cylinder, and furthermore it depends on pressure and temperature.

The calibration of a pressure balance involves taking measurements p_i , m_i , t_i of pressure, applied load and temperature, and determining from these the values of parameters A_0 and λ that define the effective area $A(p)$ as a function of pressure:

$$A(p) = A_0(1 + \lambda p).$$

In this report we discuss and compare various approaches to analysing these measurements that provide estimates of A_0 and λ . The basis of the approaches considered here is the *pressure equation* (see Section 3). The approaches differ (a) in the way that the pressure equation is used to define a data fitting problem whose solution gives estimates of A_0 and λ , and (b) in the information required about the pressure balance obtained from a prior characterization of the balance.

Depending on the nature of the measurement errors in the data, we may use the pressure equation directly to define two estimators to obtain values for the effective area parameters A_0 and λ . If there are measurement errors in the values of applied load only, we formulate a linear least-squares problem; if there are measurement errors in both the values of pressure and applied load, it is appropriate to solve a (nonlinear) generalized distance regression problem. Given knowledge of the relative sizes of the data errors, we expect these formulations to return unbiased and efficient estimates (see Section 2) of the parameters. In this sense, we regard these estimators as “benchmarks” against which other approaches may be assessed.

In practice, two methods, the so-called P - and ΔP -methods, are used to compute estimates of A_0 and λ . The P -method comes from a rearrangement of the pressure equation motivated by the experimental procedure of *cross-floating* used in a calibration. The method can be shown to correspond to a “weighted” version of the linear least-squares problem indicated above. The method is appropriate if the measurement errors occur in the values of applied load and if the standard deviations of these errors are proportional to pressure.

The ΔP -method is motivated by a relationship between *differences* in pressure and *differences* in applied load in which terms requiring a prior characterization of the balance are eliminated. By relating the method to the pressure equation, we find that the ΔP -method is appropriate if p_i , $i \geq 1$, and m_1 are exact, and if the errors in m_i , $i > 1$, are proportional to $p_i - p_1$.

The results summarised here indicate the circumstances under which it is appropriate to use the P - and ΔP -methods. If these circumstances do not apply, the methods can be expected to produce biased estimates of the effective area parameters that are less accurate than those returned by a least-squares or generalized regression method appropriate to the given circumstances. Furthermore, the assumptions necessary to ensure the appropriateness of the ΔP -method appear (a) to be more restrictive, and (b) to depart further from what can be assumed in practice than those for the P -method. In this work, we aim to justify and, in particular, *quantify* these observations by assessing the performance of the various methods on a variety

of simulated data sets

Test data sets are used to verify numerically the properties of the estimators described above, and to assess how well these estimators are able to approximate the effective area parameters for a pressure balance. Data sets are generated (a) to satisfy exactly the pressure equation, (b) to have random measurement error in the values of applied load, and optionally in the values of pressure, and (c) to have systematic error in the values of applied load as well as random error in the data values.

In all cases, the aim is to generate data that is typical of a calibration (see Section 5.4). Consequently, the numerical simulations performed used data sets constructed to have pressure values that ranged from 10 to 120MPa giving a measurement range of 12 to 1. Furthermore, for many of the data sets, no data was specified between 10 and 30MPa, *ie* within, approximately, the first 20% of the measurement range, which is a mandatory omission in the practical use of the ΔP -method. The size of the *random* mass errors (similarly pressure errors) measured as a fraction of the values of applied load (similarly pressure) corresponded to errors on average of approximately 1 part-per-million (ppm), a value that is likely to be exceeded in many laboratories. Furthermore, the size of the *systematic* error added to the values of applied load corresponded to an error of 1 gramme.

Under these circumstances, we note the following results:

1. For a large number of data sets constructed to have the properties stated above with measurement error in both applied load and pressure, and systematic error added to a subset of the mass values, the ΔP -method estimated the effective area over the complete range of pressure with an error that was on average 4.9ppm, and at worst 13.3ppm. The P -method estimated the effective area with an error that was at worst 5.3ppm, overall between two and three times better than the ΔP -method.
2. The performance of the ΔP -method is sensitive to the distribution of measured pressures. For a large number of data sets constructed as above but with measurement error in the values of applied load only, the ΔP -method estimated the effective area over the complete range of pressure with an error that was on average 1.5ppm, and at worst 5.2ppm. For data constructed to have a uniform distribution of pressure values, *ie* to include data within the first 20% of the measurement range, the error was on average 3.1ppm, and could be as large as 10.8ppm. Furthermore, for data constructed to have a uniform distribution of pressure values between 10 and 100MPa, *ie* over a smaller measurement range, the error was on average 4.1ppm, and at worst 14.2ppm.

In many laboratories, particularly at the accredited level, the random errors associated with cross-floating far exceed the 1ppm quoted above. In this case the differences between the results returned by the P and ΔP -methods can be expected to be greater than as indicated above and in the following sections.

Our more general conclusions are that:

Methods that require prior knowledge about the pressure balance obtained from a characterization of the balance perform consistently better than those not requiring such information. This suggests there are definite benefits in providing such prior knowledge.

2. The ΔP -method introduces a strongly unequal weighting to the observation equations and, unless the accuracies of the measurements justify this weighting, the method will

provide, on average, less accurate estimates of the effective area than those provided by the P -method.

3. For the types of data considered here (which we believe to be representative of the measurement system), the P - and ΔP -methods give estimates of the effective area parameters which differ by amounts that can be large compared with required accuracies. This has potentially serious consequences in terms of maintaining traceability to a primary standard.

2 INTRODUCTION

The overall goal of measurement data analysis is to provide accurate estimates of the parameters of the underlying system which gave rise to the data. The main components of the estimation process are (a) a mathematical model which specifies the system in terms of parameters $\mathbf{a}$, (b) measurement data X generated by a system specified by unknown parameter values $\mathbf{a}^*$, and (c) an estimation method or *estimator* which computes the values $\hat{\mathbf{a}}$ of parameters which best-fits the data X in some well-defined sense.

A good estimator should have the following properties: (a) *lack of bias*: if the experiment is repeated many times, the average values of the estimates should approach the true values, and (b) *efficiency*: if the experiment is repeated many times, the spread of values for the estimates should be small. Taken together these properties imply that estimates provided by one experiment should be close to the true values $\mathbf{a}^*$ of the parameters.

To define a good estimator requires (a) a valid mathematical model of the underlying system, and (b) knowledge about the type and distribution of the measurement errors that are likely to occur in the data X .

A simple example is provided by a system where a single response variable y depends on a single independent variable x linearly in terms of parameters a and b :

$$y = a + bx$$

Suppose the x -variable is set accurately at values x_i and the corresponding responses y_i are measured subject to measurement error e_i so that

$$y_i = a^* + b^*x_i + e_i,$$

for some unknown parameter values a^* and b^* .

If we assume that the measurement errors e_i represent a sampling from a normal (Gaussian) probability distribution, it can be shown (see, for example, [6]) that the best estimator or BLUE – *best linear unbiased estimator* – of (a^*, b^*) is the least-squares estimator that calculates the $\hat{a}$ and $\hat{b}$ which minimise

$$\sum_i e_i^2 = \sum_i (y_i - a - bx_i)^2$$

with respect to a and b . In other words, of all the methods of analysing this type of data the least-squares estimator will provide on average the most accurate estimates of (a^*, b^*) .

For the univariate regression example above a best estimator can be found. In practice, for more complicated models and measurements procedures, it is unlikely that an experimental situation can be completely characterized so that a best estimator can be defined. (Even if a best estimator can be defined it may be difficult to compute.) The task then becomes to use the information that is available in order to define an estimator that is computable and reasonably efficient.

In this report, however, we are not so much concerned with the determination of best estimators as in analysing two estimators already in use in pressure balance calibration. We examine under what experimental conditions they are close to optimal and use numerical simulation (based on actual experiments) to analyse their behaviour for a number of different measurement error scenarios. In Section 3 we define the mathematical models which govern the behaviour of an ideal pressure balance and discuss briefly the type of experiment that is performed in balance calibration. In Section 4 we introduce the two estimators of interest, the P - and ΔP -methods

and show how they relate to least-squares estimators. In Section 5 our data generation strategies are discussed and in Section 6 we record and analyse the results of our numerical simulations. Section 7 contains our concluding remarks. An appendix lists software implementations of the various estimators considered.

3 MODELS

3.1 THE PRESSURE BALANCE AS A PRESSURE GENERATOR

At temperature t with applied mass m (corrected for air buoyancy), a pressure balance generates a pressure p given implicitly by

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

where ϕ is a known function of t that accounts for a temperature correction, c is a known (measured) constant obtained from a precise characterization of the balance, $A(p, \mathbf{a})$ describes the effective area behaviour of the balance in terms of the pressure p and calibration parameters $\mathbf{a}$, and g is the gravitational constant.

In practice, (a) $\phi(t) = \alpha(t - 20)$ for some known α , and is small in magnitude relative to unity, and (b) A is approximately linear. Almost always A is modelled as

$$A(p, \mathbf{a}) = a_1 + a_2 p,$$

with $a_2 p$ small compared to a_1 . Substituting for $\phi(t)$ and $A(p, \mathbf{a})$, we see that (1) is equivalent to the usual form of the pressure equation, viz

$$p = \frac{(m + c)g}{A_0(1 + \lambda p)(1 + \alpha(t - 20))},$$

found, for example, in [5]; the effective area parameters A_0 and λ are related to the parameters $\mathbf{a}$ using

$$A_0 = a_1$$

and

$$\lambda = a_2/a_1$$

3.2 CALIBRATION OF A PRESSURE BALANCE

Given a set of measurements p_i, m_i, t_i of pressure, applied load and temperature, and knowledge of c , calibrating a pressure balance means finding values for the parameters $\mathbf{a}$ that best-fit the model equations

$$p_i A(p_i, \mathbf{a}) \approx \frac{(m_i + c)g}{1 + \phi(t_i)}.$$

If the values m_i are subject to *random* measurement error but the other terms are known accurately, the best-fit $\mathbf{a}$ can be determined by solving the standard least-squares problem

$$\min_{\mathbf{a}} \sum_i w_i^2 \delta_i^2, \quad (3)$$

where

$$p_i A(p_i, \mathbf{a}) = \frac{(m_i + \delta_i + c)g}{1 + \phi(t_i)}.$$

Here, (4) defines δ_i in terms of the other quantities and, on substituting these expressions for δ_i into (3), gives, if $A(p, \mathbf{a})$ is linear in $\mathbf{a}$, a linear problem for $\mathbf{a}$. The “weights” w_i in this formulation are chosen to be inversely proportional to the standard deviations of the errors in the masses m_i . The solution can be shown to be unbiased and, of all linear estimators, to be the estimator of minimum standard error. In other words, the estimator defined by (3) and (4) is the best linear unbiased estimator (BLUE) for a system described by the pressure equation (1) using data with measurement error in the values of applied load.

However, if the values p_i as well as m_i are subject to random measurement error, it is appropriate to use a generalized regression approach in which the optimal $\mathbf{a}$ is determined by solving

$$\min_{\mathbf{a}, \epsilon, \delta} \sum_i \alpha_i^2 \epsilon_i^2 + \beta_i^2 \delta_i^2 \quad (5)$$

subject to

$$(p_i + \epsilon_i)A(p_i + \epsilon_i, \mathbf{a}) = \frac{(m_i + \delta_i + c)g}{1 + \phi(t_i)}.$$

This can be solved efficiently with standard optimisation software using a generalized distance regression approach [1]. Again, estimates of the likely sizes of the errors in the measured values of pressure and mass are necessary in order to assign the weights α_i and β_i .

In a cross-float experiment [5, 8] the pressure p is generated by connecting the test balance to a reference balance with known effective area parameters $\mathbf{b}$ and adjusting the loads on the two balances so that both are in equilibrium. Suppose M is the load applied to the reference balance, C is a known correction to this load obtained from a characterization of the balance, and $\Phi(T)$ is a temperature correction. Then, when the balances are in equilibrium, we have

$$p = \frac{(M + C)g}{A(p, \mathbf{b})(1 + \Phi(T))} = \frac{(m + c)g}{A(p, \mathbf{a})(1 + \phi(t))}. \quad (6)$$

The pressure p is determined from the reference balance parameters $\mathbf{b}$, and its value will be wrong if the parameters are wrong. However, these parameters are perfectly convolved with the unknowns $\mathbf{a}$, so the only option is to accept $\mathbf{b}$ as accurate. Estimates of their uncertainty can be fed through in providing uncertainties for $\mathbf{a}$.

Finally, we note the possibility that there may be *systematic* errors in the masses m_i due, for example, to operator error. Typically, the k th applied load m_k may be misread and recorded as $m_k + \gamma$, where γ is an unknown systematic error in the applied load, and m_k may contain measurement error. In this case, because of the incremental way that masses are added, the values recorded for m_i , $i > k$, may be contaminated in a similar way.

4 ESTIMATORS

In the following descriptions we assume, as in Section 3, that the effective area is modelled as

$$A(p, \mathbf{a}) = a_1 + a_2 p.$$

4.1 LINEAR LEAST-SQUARES METHOD

The least-squares problem defined by (3) and (4) may be written as

$$\min_{\mathbf{a}} \sum_i w_i^2 \delta_i^2,$$

where

$$(a_1 p_i + a_2 p_i^2)(1 + \phi(t_i)) = (m_i + \delta_i + c)g.$$

As discussed above, this approach gives the best linear unbiased estimates of the effective area parameters if the measurement errors occur only in the values m_i of mass. In Appendix B.1 we provide a listing of a Matlab [7] function **wllspe** for solving this problem. The function returns estimates of the effective area parameters A_0 and λ , together with the values of applied load that are predicted at the values of pressure p_i by the pressure equation (1) using these estimates. (We describe in Section 5 how we use the Matlab module **pmass** to compute these values of applied load, and we discuss also their significance.)

We also provide a listing of a Matlab function **wllspec** that solves the above least-squares problem but regarding c , in addition to the parameters $\mathbf{a}$, as unknown. We include this function because, like the ΔP -method considered below, values of the effective area parameters are returned without the need to specify c from a prior characterization of the pressure balance. Observe that with c as an unknown, the least-squares problem formulated above remains a *linear* problem.

4.2 P-METHOD

The right hand side equality in (6) can be rearranged to read

$$\frac{a_1 + a_2 p}{b_1 + b_2 p} = \frac{m + c}{M + C} \frac{1 + \Phi(T)}{1 + \phi(t)},$$

from which we obtain the model equations

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

where $\mathbf{b}$, m_i , M_i , t_i , T_i , c , C and p_i are known or measured. The P -method consists of solving the equations (7) in a least-squares sense for the parameters $\mathbf{a}$.

The model is essentially equivalent to

$$\left\{ a_1 p + a_2 p^2 = \frac{(m + c)g}{1 + \phi(t)} \right\} \times \frac{1}{p}.$$

In other words, the method is almost identical to solving the least-squares problem defined by (3) and (4) using for the weights w_i the values $1/p_i$. This gives the best linear unbiased estimates of the effective area parameters if the measurement errors occur in the values m_i and if the standard deviations of these errors are proportional to pressure.

In the usual implementation of the P -method, the left hand side of (7) is further approximated by

$$\frac{a_1 + a_2 p_i}{b_1 + b_2 p_i} \approx \frac{a_1}{b_1} \left(1 + \left(\frac{a_2}{a_1} - \frac{b_2}{b_1} \right) p_i \right) \quad (9)$$

(using the fact that $b_2 p_i$ is small compared with b_1 : cf Section 3), which reduces finding the solution of (7) to a straight-line fitting problem.

In Appendix B.2 we provide listings of Matlab functions **pmethod0** and **pmethod1** that implement the P -method, respectively, with and without the approximation (9).

4.3 ΔP -METHOD

The ΔP -method provides an approximation to $p - p_1$, viz

$$p - p_1 = \frac{(m - m_1)g}{(a_1 + a_2(p + p_1))(1 + \phi(t))} + (\phi(t_1) - \phi(t))p_1,$$

chosen in such a way as to eliminate c from the pressure equation [4]. Rearranging we have

$$a_1 + a_2(p + p_1) = \frac{(m - m_1)g}{(p - p_1 + (\phi(t) - \phi(t_1))p_1)(1 + \phi(t))}, \quad (10)$$

from which we obtain the model equations

$$a_1 + a_2(p_i + p_1) = \frac{(m_i - m_1)g}{(p_i - p_1 + (\phi(t_i) - \phi(t_1))p_1)(1 + \phi(t_i))} + e_i, \quad i > 1, \quad (11)$$

where m_i , p_i and t_i are known or measured. The ΔP -method consists of solving the equations (11) in a least-squares sense for the parameters $\mathbf{a}$.

We note that the left hand side of (10) is

$$(a_1p + a_2p^2 - a_1p_1 - a_2p_1^2) \times \frac{1}{p - p_1} \quad (12)$$

and the right hand side is an approximation to

$$\left\{ \frac{(m + c)g}{1 + \phi(t)} - \frac{(m_1 + c)g}{1 + \phi(t_1)} \right\} \times \frac{1}{p - p_1}. \quad (13)$$

the approximation being accurate if $\phi(t) = \phi(t_1)$. Consequently, the method essentially reduces to finding a solution to the least-squares problem defined by (3) and (4) that satisfies exactly the pressure equation evaluated using m_1 , p_1 and t_1 , and with the weights w_i assigned the values $1/(p_i - p_1)$, $i > 1$. This would give the best linear unbiased estimates of the effective area parameters if p_i , $i \geq 1$, and m_1 are exact, and if the errors in m_i , $i > 1$ are proportional to $p_i - p_1$.

In Appendix B.3 we provide listings of Matlab functions **dpmethod0** and **dpmethod1**. The first of these implements the ΔP -method using (11) and does not require us to supply a value for c . The second function implements the estimator using (13) in place of the right hand side in (11). In Section 6 we use these functions to show that the two ways described above of looking at the ΔP -method are indeed equivalent.

We also give a listing of the Matlab function **dpmethodi** that implements the ΔP -method but using the data m_{i_0} , p_{i_0} and t_{i_0} from the i_0 th measurement, rather than the first measurement, to eliminate c from the pressure equation. This is expected to have the effect of shifting the bias of the estimator towards the i_0 th point and its neighbouring measurements, and this is investigated numerically in Section 6.

4.4 MEASUREMENT STRATEGY

We have seen that the P and ΔP -methods can be viewed as least-squares estimators with weights proportional to $1/p_i$ and $1/(p_i - p_1)$, respectively. In practice, their efficiency is therefore related to the extent to which the weights reflect actual measurement error. Moreover, the

measurement strategy can be chosen to minimise biasing effects. Suppose, for example, that it is known that the measurement errors in applied load are independent of pressure and have equal variance. Then, measuring at pressures 30, 40, 50, ..., 120MPa is better than measuring at 10, 20, 30, ..., 100MPa for the P -method since the weighting $1/p_i$ is closer to being uniform for the first strategy. Similarly, measuring at pressures 10, 40, 50, ..., 120MPa is better than measuring at 10, 20, 30, ..., 100MPa if the ΔP -method is used for the same reason. A major difference between the two methods is that the ΔP -method treats one measurement as exact (which can be approximated using a very large weighting) whereas the P -method models measurement error in all measurements.

5 DATA GENERATION

In this section we describe our approach to generating test data sets. The data sets are used (a) to verify numerically the properties of the various estimators outlined in Section 4, and (b) to assess how well these estimators are able to approximate the effective area for a pressure balance. The results of numerical experiments using the generated test data sets are presented in Section 6.

The starting point for constructing test data is, given values for the effective area parameters, to generate data that satisfies exactly the pressure equation (1) or its equivalent form (2). Given such data we then add random measurement error, and optionally systematic error, in order to simulate data generated by a real experiment.

Random measurement error can be added in two ways. Firstly, we can add to the values m_i , and optionally to the values p_i , samples from a Gaussian probability distribution function that has zero mean and a standard deviation determined by our knowledge of the likely accuracy of these quantities. In this way we can assess and compare how well the various estimators described in Section 4 recover the values of the effective area parameters for the test balance.

Secondly, we can add perturbations to the values m_i , and optionally to the values p_i , that are chosen so that a specified estimator, such as the standard least-squares or generalized regression estimator, returns the given values of the effective area parameters. In this way we can compare the P - and ΔP -methods with what we believe to be the *optimal* estimator for those data sets. An important advantage of this approach to generating test data sets is that it is not necessary for us to implement the specified estimator because, by construction, we know *a priori* the solution the estimator will return. This is particularly useful for the generalized distance regression estimator for which we have provided no implementation.

5.1 GENERATING DATA TO SATISFY EXACTLY THE PRESSURE EQUATION

As a first stage in the process of data generation we use equation (2) to generate a set of self-consistent data for a given balance in the following way. We prescribe values to the balance parameters A_0 , λ , α and c , and choose a set of pressure values p_i , $i = 1, \dots, n$, and temperature values t_i , $i = 1, \dots, n$. A rearrangement of (2) is then used to evaluate the masses m_i , $i = 1, \dots, n$, that generate those pressures for the given balance. We apply this procedure to generate data relating to two different balances, the *test* and *reference* balances, using a common set of pressure values but perhaps different temperature values for each balance.

In Appendix B.4 we give listings of Matlab functions **pconst** and **pmass**. Function **pconst** is used to assign values for the balance parameters used in the numerical experiments described in

Section 6. Function **pmass** can then be used to apply the pressure equation to give values m_i that are consistent with the balance parameters and chosen values of pressure and temperature.

5.2 LEAST-SQUARES DATA GENERATOR

Consider the standard linear least-squares (LLS) problem

$$\min_{\mathbf{a}} \sum_{i \in I} w_i^2 (y_i - f(x_i, \mathbf{a}))^2, \quad (14)$$

where the function f is linear in the unknown parameters $\mathbf{a}$. (The strategy described below also applies, with minor modification, to nonlinear functions.) By equating the gradient of this sum to zero, it is easy to derive conditions on $\mathbf{a}$ which are necessary for a local minimum, *viz*

$$\sum_{i \in I} w_i^2 (y_i - f_i) \nabla_{\mathbf{a}} f_i = \mathbf{0},$$

where $f_i = f(x_i, \mathbf{a})$. These conditions lead naturally to a simple method [2] for generating data points (x_i, y_i) for which the solution to the corresponding LLS problem (14) is known *a priori*. Given x_i , w_i and $\mathbf{a}$,

I Calculate $f_i = f(x_i, \mathbf{a})$

II Calculate the matrix X with elements

$$X_{i,j} = \frac{\partial f_i}{\partial a_j}$$

III Determine any vector $\mathbf{t}$ such that

$$X^T \mathbf{t} = \mathbf{0}.$$

IV Set

$$y_i = f_i + t_i / w_i^2.$$

It is easily verified that with this choice of points (x_i, y_i) , the conditions for a solution are satisfied.

The key step in the procedure described above is the determination of a vector $\mathbf{t}$ in the null space of the matrix X^T . If X is an $n \times q$ matrix, the null space of X^T is an $(n - q)$ -dimensional subspace of n -space. Using an orthogonal factorisation of X , it is possible to find an orthonormal basis $\mathbf{z}_k$, $k = 1, \dots, n - q$, for this subspace [3]. Any vector $\mathbf{t}$ in the null space can be written as a linear combination

$$\mathbf{t} = \sum_{k=1}^{n-q} \nu_k \mathbf{z}_k \quad (15)$$

of the basis vectors $\mathbf{z}_k$ and, conversely, any choice of coefficients ν_k gives rise to a vector in the null space via equation (15). Thus, by generating sets of coefficients ν_k (using a random number generator, for example) and following steps I to IV, it is possible to generate an arbitrary number of data sets for which the least-squares solution is $\mathbf{a}$.

5.3 GENERALIZED DISTANCE REGRESSION DATA GENERATOR

Consider the generalized distance regression (GDR) problem

$$\min_{\mathbf{x}^*, \mathbf{a}} \sum_{i \in I} \{ \alpha_i^2 (x_i - x_i^*)^2 + \beta_i^2 (y_i - f(x_i^*, \mathbf{a}))^2 \}. \quad (16)$$

This is a nonlinear least-squares problem for which we can derive the following first order conditions on the parameters x_i^* , $i \in I$, and $\mathbf{a}$ which are necessary for a local minimum:

$$\alpha_i^2 (x_i - x_i^*) + \beta_i^2 (y_i - f_i) \dot{f}_i = 0,$$

and

$$\sum_{i \in I} \beta_i^2 (y_i - f_i) \nabla_{\mathbf{a}} f_i = \mathbf{0},$$

where $f_i = f_i(x_i^*, \mathbf{a})$ and $\dot{f}_i = \frac{\partial f_i}{\partial x}$. These conditions lead to a method for generating data points (x_i, y_i) for which the solution of the corresponding GDR problem (16) is known *a priori*. Given x_i^* , α_i , β_i and $\mathbf{a}$,

I Calculate $f_i = f(x_i^*, \mathbf{a})$.

II Calculate $\dot{f}_i = \frac{\partial f_i}{\partial x}$.

III Calculate the matrix X with elements

$$X_{i,j} = \frac{\partial f_i}{\partial a_j}$$

IV Determine any vector $\mathbf{t}$ such that

$$X^T \mathbf{t} = \mathbf{0}$$

V Set

$$\begin{aligned} x_i &= x_i^* - t_i \dot{f}_i / \alpha_i^2, \\ y_i &= f_i + t_i / \beta_i^2. \end{aligned}$$

It is easily verified that with this choice of points (x_i, y_i) , the conditions for a solution are satisfied. Moreover, as for the LLS problem, it is possible to use this procedure to generate an arbitrary number of data sets with the same GDR solution. The Matlab module **ndgsgdr.m**, the specification of which is given in Appendix B.5 is a generic function that executes steps I to V above, given a module **fgeval.m** which calculates the information specific to f .

The GDR problem for calibrating a pressure balance can be written as

$$\min_{a_1, a_2, \mathbf{p}^*} \{ \alpha_i^2 (p_i - p_i^*)^2 + \beta_i^2 (m_i - f_i)^2 \},$$

where

$$f_i = (1 + \phi(t_i)) p_i^* (a_1 + a_2 p_i^*) / g - c.$$

This is equivalent to (5) with

$$p_i^* = p_i + \epsilon_i$$

and

$$f_i = m_i + \delta_i.$$

Data sets for the GDR problem can be generated using **ndgsgdr.m** in conjunction with the module **pbfggeval.m** listed in Appendix B.5. The function **pbfggeval.m** evaluates f_i (as defined above) together with the derivatives of f_i with respect to pressure p and the effective area parameters $\mathbf{a}$.

5.4 GENERATING TYPICAL DATA

For the results of any numerical simulation to be meaningful it is important that the data sets used in the simulation are typical of real experiments. In this section we record the major decisions made affecting the generation of test data for the numerical experiments described in Section 6.

Parameters for the test and reference balances. The Matlab function **pconst** is used to assign values to the parameters for the test and reference balances, including those parameters to be estimated. A listing of this function is given in Appendix B.4.

Values of temperature. For most of the numerical experiments, the values t_i and T_i of temperature for, respectively, the test and reference balances are set equal to $20 + \omega_i$ ($^{\circ}\text{C}$) and $20 + \Omega_i$ ($^{\circ}\text{C}$), where ω_i and Ω_i are samples taken from a Gaussian probability distribution with mean zero. (Some numerical experiments are included for which the temperatures t_i and T_i are each assigned to be constant: see Section 6.)

Values of pressure. Unless otherwise stated, the (nominal) values p_i of pressure at which data is generated are

$$10, 30, 35, 40, 45, 50, \dots, 110, 115, 120\text{MPa}.$$

This gives a measurement range of 12 to 1 and, furthermore, the distribution of pressure values is such that no data is present within, approximately, the first 20% of the measurement range. These properties of the distribution of pressure values are typical in the use of the ΔP -method as it mollifies the implicit “ $1/(p_i - p_1)$ ” weighting effect associated with the method. For this distribution of pressure values, the relative weighting associated with the ΔP -method is 5.5 to 1, whereas for the P -method it is 12 to 1. (Some numerical experiments are included for which the values of pressure are uniformly spaced throughout the given range as well as uniformly spaced throughout a smaller measurement range: see Section 6.)

Values of applied load. Given values of pressure, temperature and the parameters describing a pressure balance, values for the applied load m_i that satisfy exactly the pressure equation are generated using the Matlab function **pmass** as described in Section 5.1. A listing of this function is given in Appendix B.4.

Random measurement error for values of applied load only. An example set of mass errors, generated in such a way that the least-squares estimator returns the true values of the effective area parameters, is listed in Table 4. We present in this table the errors as a difference between the assigned data value m_i and the true value m_i^* , viz

$$m_i - m_i^*,$$

measured in milligrammes (mg), and as a proportion of the values m_i^* to which they are applied, viz

$$\frac{|m_i - m_i^*|}{m_i^*} \times 10^6,$$

measured in parts-per-million (ppm).

The sizes of the mass errors are controlled by requiring that the quantity

$$\frac{1}{n} \sqrt{\sum_{i=1}^n (m_i - m_i^*)^2},$$

that measures the discrepancy between the data and the true loads, has a value of 25mg for a data set consisting of $n = 20$ points (as in the distribution of pressures listed above). In terms of parts-per-million, the mass errors are largest for the data generated at low pressures for which the mass values m_i^* are smallest. However, across the complete pressure range, the average mass error is typically approximately 1ppm.

Random measurement error for values of applied load and pressure. An example set of errors for applied load and pressure, generated in such a way that the generalized distance regression estimator returns the true values of the effective area parameters, is listed in Table 10. We list in this table the errors as a difference between the assigned data value and its true value, and as a proportion of the true value (see above). The average error in applied load and pressure is typically approximately 1ppm.

Systematic error for values of applied load. A systematic error of 1 gramme is added to a subset of the values of applied load for the test balance.

5.5 COMPARISON OF RESULTS

For a single data set, a natural way to assess the performance of a particular estimator is to compare the values for the effective area parameters returned by the estimator with the effective area parameters used in the construction of the data set.

If many different, but similarly constructed, data sets are available, it is useful to consider the distributions of the results returned by the estimator. The *arithmetic mean* of the distribution of values for A_0 (and similarly for λ) tells us how the estimator behaves “on average” by highlighting *bias* in the results if such bias exists. The *standard deviations* of these distributions indicates the average departure of the results from the mean, and so measures how accurate individual results can be expected to be. Furthermore, the *range* of the distributions, as determined by the minimum and maximum computed values, provides a *worst-case* measure for the performance of the estimator.

Looking at the distributions of the individual effective area parameters A_0 and λ is a natural way to assess the performance of the methods considered, but a measure of perhaps more practical interest is the error in the effective area $A(p, \mathbf{a})$ itself. If $\mathbf{a}^*$ and $\mathbf{a}$ denote, respectively, the given and estimated effective area parameters, the quantity

$$E_A = \max_{p_i} |A(p_i, \mathbf{a}^*) - A(p_i, \mathbf{a})|$$

provides a single measure of the discrepancy in the estimated effective area over the range of the data. In the numerical simulations described in Section 6 we often measure E_A in parts-per-million (ppm) by giving the value of the quantity

$$\frac{E_A}{A_0} \times 10^6.$$

Finally, we can think of the relationship between pressure and applied load as defining a *calibration curve* for a given pressure balance. Having computed estimates of the effective area parameters for the balance, we apply the pressure equation (1) or (2) with these estimates to predict values $\hat{m}_i$ for the true applied loads m_i^* that generate the given pressures p_i . The differences between $\hat{m}_i$ and m_i^* give us an indication of how well the calibration curve is being estimated. The values $\hat{m}_i - m_i^*$ are the errors in applied load that result from the errors in the estimates of the effective area parameters. Clearly, we would hope that the predicted loads $\hat{m}_i$

are at least as close to the true values m_i^* as the data values m_i . Furthermore, *trends* in the values $\hat{m}_i - m_i^*$ indicate a bias in the results that may not be apparent from looking at the estimates of the effective area parameters. The quantity

$$E_m = \frac{1}{n} \sqrt{\sum_{i=1}^n (\hat{m}_i - m_i^*)^2}$$

provides a single (averaged) measure of this discrepancy which we quote in the numerical simulations carried out in Section 6. We would hope to obtain a value for E_m that is not significantly larger than the quantity

$$\frac{1}{n} \sqrt{\sum_{i=1}^n (m_i - m_i^*)^2}$$

(see Section 5.4) that measures the discrepancy between the *data* and the true applied loads.

6 NUMERICAL SIMULATIONS

6.1 EXACT DATA

Data satisfying exactly the pressure equation was generated as described in Section 5.1 using the Matlab functions **pconst** and **pmass**. Three data sets were generated:

- For the first the temperature values t_i and T_i for the test and reference balances, respectively, were each kept constant.
- For the second the temperatures were assigned values randomly distributed about a nominal temperature of 20°C , the deviations of temperature from this nominal value being samples from a Gaussian probability distribution with mean zero and standard deviation of 0.1°C .
- For the third the temperatures were assigned values randomly distributed about a nominal temperature of 20°C , the deviations of temperature from this nominal value being samples from a Gaussian probability distribution with mean zero and standard deviation of 1°C .

Tables 1, 2 and 3 show for the three data sets the results returned by the approaches described in Section 4 (here referenced by the names of their Matlab implementations: see Section 4). For the application of the least-squares estimators **wllspe** and **wllspec**, the weights w_i were set equal to unity.

Note:

1. The P -method implemented in **pmethod0** does not return the (true) effective area parameters used to generate the data. This is a consequence of using the approximation (9): compare the results for **pmethod0** and **pmethod1**.
2. The ΔP -method implemented in **dpmethod0** does not return the true effective area parameters for the data sets for which the temperatures for the measurements vary. This confirms that the derivation of (11) employs approximations that are only accurate when the temperature values are constant: compare the results for **dpmethod0** and

dpmethod1. Furthermore, the deviation of the estimated effective area parameters from their true values becomes larger as the deviation of temperature from a constant increases.

3. Nevertheless, in terms of estimating the effective area over the complete range of pressure values, the values of E_A indicate that the accuracy of the results is acceptable. (A value for E_A of $2.7 \times 10^{-7} \text{ mm}^2$ corresponds to an error of approximately 0.01ppm in the estimation of effective area.)
4. The least-squares approach **wllspec** that estimates the correction term c as well as the effective area parameters, returns the true values for these parameters for the data sets.

6.2 SIMULATING THE P - AND ΔP -METHODS

In this section we present results that illustrate the equivalence of the P -method (7) with the weighted least-squares estimator (8), and the ΔP -method (11) with the weighted least-squares estimator defined by (12) and (13).

We use two data sets for which the values of applied load for the test balance have been contaminated with measurement error in such a way that the least-squares estimator **wllspe** (applied with unit weights) returns the effective area parameters used to generate the data. Again, the data sets differ in the way the values t_i and T_i are assigned: for the first data set, the values of temperature for each balance are constant, and for the second set, the measured temperature is allowed to vary.

The mass errors for the first of the data sets are listed in Table 4, and similar errors are used to generate the second set of data. Tables 5 and 6 show the results for the various approaches considered.

Note:

As expected, the application of **wllspe** with unit weights returns the true effective area parameters.

2. For both sets of data, there is general agreement between the results from the application of **wllspe** with weights w_i set to be $1/p_i$ and those returned by **pmethod1**. The application of **wllspe** with these weights does not mimic as accurately the use of **pmethod0** because of the effect of the approximation (9) used in the P -method. (The slight discrepancy can also be attributed to the small effect of the term $1 + \phi(t_i)$.)
3. There is agreement between the results returned by **dpmethod0** and **dpmethod1**. Furthermore, the results are particularly close for the data set for which the temperature values are kept constant: this reflects the dependence of the approximation (11) used in the ΔP -method on properties of the temperature distribution.

It is evident from all the measures given that for these data sets the results returned by the P -method provide a more accurate calibration of the test balance than those returned by the ΔP -method.

6.3 THE EFFECT OF USING DIFFERENT PRIOR KNOWLEDGE IN THE ΔP -METHOD

As we noted in Section 4, the ΔP -method uses information gathered from the first measurement as prior knowledge to eliminate the correction term c , and thus remove the necessity to measure this quantity for the test balance. Of course, c can be eliminated using the information gathered from *any* of the measurements, and the Matlab function **dpmethodi** implements this generalization. Given the index i_0 , **dpmethodi** determines effective area parameters that solve the weighted least-squares problem

$$\left\{ a_1 p_i + a_2 p_i^2 - a_1 p_{i_0} - a_2 p_{i_0}^2 = \frac{(m_i + c)g}{1 + \phi(t_i)} - \frac{(m_{i_0} + c)g}{1 + \phi(t_{i_0})} \right\} \times \frac{1}{p_i - p_{i_0}}$$

(cf equations (12) and (13)).

We use a data set for which the temperature distributions are variable, and the values of applied load for the test balance are contaminated with mass errors that are similar to those listed in Table 4. For each index $i_0 = 1, \dots, 20$, we use **dpmethodi** to compute estimates of the effective area parameters. The results are summarised in Table 7.

Note:

1. Different estimates of the effective area parameters are returned for different choices of the index i_0 defining the prior knowledge used.
2. The accuracy of the results returned depends on how i_0 is chosen in relation to the sizes of the mass errors. For this data set, the accuracy of the results is worst for i_0 near 1 and 20 (corresponding to the first and last points where the data errors are largest: see Table 4). This is not an “end-effect”: if there is a large mass error near the centre of the range, choosing i_0 to coincide with this mass error will tend to give poorer results than for other choices of i_0 . (There is a minor end-effect caused by the different distributions of weights.)
3. Depending on the choice of the index i_0 defining the prior knowledge used, the error in estimating the effective area of the balance ranges from 0.3ppm to approximately 24ppm.

6.4 DATA WITH RANDOM MEASUREMENT ERROR

In this section we compare the performance of the various estimators considered in Section 4 on data constructed to include random measurement error. We consider data for which (a) the values m_i of applied load are contaminated with normally distributed measurement error, (b) the least-squares estimator returns the true parameters, and (c) the generalized distance regression estimator returns the true parameters. Rather than considering individual data sets, the performance of the estimators is assessed by comparing the distributions of the measures E_m and E_A for each estimator applied to a large number of different data sets.

6.4.1 Results for data with normally distributed random error

In this section we compare the various estimators considered using data generated by adding normally distributed random error to the values m_i of applied load for the test balance. For

this type of data, we expect the least-squares estimator **wllspe** with unit weights to be the best linear unbiased estimator.

We apply the estimators to 500 different data sets for which the values of temperature t_i and T_i are allowed to vary. The results are summarised in Table 8, where we give the mean, standard deviation and maximum values of E_m and E_A for each approach.

Note:

1. For the estimation of effective area, the error in the results returned by the ΔP -method is on average roughly twice that for the P -method.
2. Methods that require prior knowledge about the test balance obtained from a characterization of the balance perform better than those not requiring such information. Furthermore, the results for the estimator **wllspec** indicate that its performance is competitive with the ΔP -method.

6.4.2 Results for data for which the LLS estimator returns the true parameter values

In this section we compare the various estimators considered using data generated in such a way that the least-squares estimator (with unit weights) returns the effective area parameters used to generate the data.

We apply the estimators to 500 different data sets generated to have this property. The results are summarised in Table 9.

Note:

1. As expected, the least-squares estimator **wllspe** returns the true effective area parameters.
2. For the estimation of effective area, the error in the results returned by the ΔP -method is on average roughly twice that for the P -method.
3. The data sets were generated in such a way that we expect E_m not to be significantly larger than 2.5×10^{-5} kg. For the least-squares and P -methods, this expectation is confirmed. However, the distribution of E_m values for the ΔP -method suggests that for this type of error structure in the data there may be significant bias in the values of applied load derived from the estimated effective area parameters.
4. The results presented in Table 9 are comparable with those given in Table 8 for data with normally distributed measurement error. This is not surprising since the magnitudes of applied mass errors are comparable for the two cases. In particular, the error in estimating effective area using the ΔP -method is on average 1.5ppm, and at worst approximately 5ppm.

6.4.3 Results for data for which the GDR estimator returns the true parameter values

In this section we compare the various estimators considered using data generated in such a way that the generalized distance regression estimator returns the effective area parameters used to generate the data.

We apply the estimators to 500 different data sets generated to have the property that the GDR estimator for which $\alpha_i = \alpha$, $\beta_i = \beta$ and $\alpha = 4\beta$ returns the given effective area parameters. The results are summarised in Table 11.

Note:

1. Introducing measurement error into the pressure values p_i as well as the values of applied load has reduced the accuracy of the results returned by all the approaches. Recall that all the approaches tested here are based on the assumption that the values p_i are known accurately.
2. For these data sets, the error in estimating the effective area using the effective area parameters returned by the ΔP -method is on average 3.1ppm and can be as large as 9.3ppm. These are roughly three times the corresponding values for the P -method.
3. The least-squares estimator **wllspec** appears to be competitive with the ΔP -method, but not with the P -method, indicating the importance of measuring the correction term, c .

6.5 DATA WITH RANDOM MEASUREMENT ERROR AND SYSTEMATIC ERROR

In this section we compare the performance of the various estimators considered in Section 4 using data constructed to include systematic error as well as random measurement error. Of all the data considered thus far, data with both types of error is probably the most realistic.

We apply the estimators to 500 different data sets generated in the following way. Firstly, random measurement error is added to the values p_i of pressure and the values m_i of applied load in such a way that the generalized distance regression estimator returns the effective area parameters used to generate the data (see Section 6.4.3). Secondly, a systematic error of 1 gramme is added to the last 10 of the 20 mass values. The results are summarised in Table 12.

Note:

1. Comparing the results for data with and without systematic error (Tables 11 and 12, the most significant change is in the mean and maximum values of the measures E_m and E_A that indicates that the systematic error introduces a bias in the results.
2. For these data sets, the error in estimating the effective area using the effective area parameters returned by the ΔP -method is on average 4.9ppm and can be as large as 13.3ppm. For the P -method the error is at worst 5.3ppm.

6.6 DATA BASED ON A UNIFORM DISTRIBUTION OF PRESSURE VALUES

We conclude by comparing the performance of the various estimators using data sets generated using: (a) a distribution of pressure values that includes data within the first 20% of the measurement range, and (b) a distribution of pressure values that define a measurement range of 10 to 1.

The data sets are generated as in Section 6.4.2 but using values p_i that are uniformly distributed throughout the following measurement ranges:

- Uniformly between 10 and 120MPa, *viz*

10, 15, 20, 25, 30, 35, ..., 110, 115, 120MPa,

giving, as before, a measurement range of 12 to 1

- Uniformly between 10 and 100MPa, *viz*

10, 15, 20, 25, 30, 35, ..., 90, 95, 100MPa,

giving a reduced measurement range of 10 to 1

The results are summarised in Tables 13 and 14, respectively.

Note:

1. Comparing the results in Table 9 with those in Table 13, it is apparent that the inclusion of additional data near the beginning of the pressure range has a positive affect on all the approaches except for the ΔP -method for which the estimation of effective area is made less accurate. For these data sets, the error in estimating the effective area using the parameters returned by the ΔP -method is on average 3.1ppm and can be as large as 10.8ppm. For the P -method the error is at worst 1.8ppm.
2. Comparing the results in Table 14 with those in Table 13, it is apparent that reducing the measurement range by removing measurements from one end of the measurement range makes the results returned by all methods less accurate. However, the absolute change in the results is greatest for the ΔP -method, although the relative change is much the same for both P and ΔP -methods. For these data sets, the error in estimating the effective area using the parameters returned by the ΔP -method is on average 4.1ppm and can be as large as 14.2ppm. For the P -method the error is at worst 2.4ppm.

7 CONCLUSIONS

In this report we have discussed and compared various existing and proposed approaches for estimating the effective area parameters for a pressure balance. We have presented the circumstances under which those approaches provide best linear unbiased estimates for the parameters, and have quantified the performance of the methods using numerical simulation. Test data sets have been obtained from model-based data generators to simulate (appropriate combinations of):

- no measurement error
- random measurement error in the values of applied load
- random measurement error in the values of pressure:
- systematic error in the values of applied load;
- uniformly and non-uniformly spaced values of pressure;
- constant and variable temperature distributions.

The performance of the various approaches has been assessed using the results obtained for individual data sets and by considering the distributions of results for large numbers of data sets. Measures of performance have included the ability to estimate the effective area (E_A), as well as the ability to approximate satisfactorily the data (E_m).

We conclude that (under the assumptions used to generate the data)

1. Methods that require prior knowledge about the pressure balance obtained from a characterization of the balance perform consistently better than those not requiring such information. This suggests there are definite benefits in providing such prior knowledge.
2. The ΔP -method is sensitive to measurement strategy. In some circumstances, the method introduces a strongly unequal weighting to the observation equations and, unless the accuracies of the measurements justify this weighting, the method will provide, on average, less accurate estimates of the effective area than those provided by the P -method.
3. For the types of data considered here (which we believe to be representative of the measurement system), the P - and ΔP -methods give estimates of the effective area parameters which differ from each other by amounts that are significant compared with required accuracies. This has potentially serious consequences in terms of maintaining traceability to a primary standard.
4. Estimators based on linear least-squares (LLS), generalized distance regression (GDR) or the P -method can be “tuned” to give efficient estimators through the use of appropriate weights. The ΔP -method cannot be as easily tuned because it treats one measurement as exact.

8 ACKNOWLEDGEMENTS

We thank our colleagues David Simpson, Phil Clow and Maurice Cox for commenting on earlier drafts of this report.

9 REFERENCES

- [1] B.P. Butler, A.B. Forbes, and P.M. Harris. Geometric tolerance assessment problems. In P. Ciarlini, M.G. Cox, R. Monaco, and F. Pavese, editors, *Advanced Mathematical Tools in Metrology*, pages 95–104. World Scientific, Singapore, 1994.
- [2] M.G. Cox and A.B. Forbes. Strategies for testing form assessment software. Technical Report DITC 211/92, National Physical Laboratory, Teddington, 1992.
- [3] G.H. Golub and C.F. Van Loan. *Matrix Computations*. North Oxford Academic, Oxford, 1983.
- [4] J.C. Legras and P. Delajoud. Comparaison des étalons de pression du Gosstandart (URSS) et du Bureau National de Métrologie (France) avec la balance de transfert de pression du Bureau National de Métrologie (Gamme 1 à 7 mpa). *Bulletin BNM*, 1977.
- [5] S. Lewis and G. Peggs. *The Pressure Balance: A Practical Guide to its Use*. HMSO London, second edition, 1992.
- [6] K.V. Mardia, J.T. Kent, and J.M. Bibby. *Multivariate Analysis*. Academic Press, London, 1979.
- [7] C. Moler, J. Little, and S. Bangert. *PRO-MATLAB: User's Guide*. The MathWorks, Inc., South Natick, MA, USA, 1988.
- [8] D.I. Simpson. Computerized techniques for calibrating pressure balances. *Metrologia*, 30:655–658, 1993.

A NUMERICAL RESULTS

	$A_0(mm^2)$	$\lambda(/MPa) \times 10^6$	$c(kg) \times 10^2$	$E_m(kg)$	$E_A(mm^2)$
wllspe	30.6	3.93		0	0
wllspec	30.6	3.93	3.7	0	0
pmethod0	30.59999992	3.93009631		3.6×10^{-7}	2.7×10^{-7}
pmethod1	30.6	3.93		0	0
dpmethod0	30.6	3.93		0	0
dpmethod1	30.6	3.93		0	0

Table 1 Results for data satisfying exactly the pressure equation: temperatures for the test and reference pressure balances are constant.

	$A_0(mm^2)$	$\lambda(/MPa) \times 10^6$	$c(kg) \times 10^2$	$E_m(kg)$	$E_A(mm^2)$
wllspe	30.6	3.93		0	0
wllspec	30.6	3.93	3.7	0	0
pmethod0	30.59999992	3.93009631		3.6×10^{-7}	2.7×10^{-7}
pmethod1	30.6	3.93		0	0
dpmethod0	30.60000000	3.93000047		5.6×10^{-9}	4.3×10^{-9}
dpmethod1	30.6	3.93		0	0

Table 2 Results for data satisfying exactly the pressure equation: deviations from 20°C of the temperatures for the test and reference pressure balances are samples from a Gaussian probability distribution with mean zero and standard deviation of 0.1°C.

	$A_0(mm^2)$	$\lambda(/MPa) \times 10^6$	$c(kg) \times 10^2$	$E_m(kg)$	$E_A(mm^2)$
wllspe	30.6	3.93		0	0
wllspec	30.6	3.93	3.7	0	0
pmethod0	30.59999992	3.93009631		3.6×10^{-7}	2.7×10^{-7}
pmethod1	30.6	3.93		0	0
dpmethod0	30.59999997	3.92999342		9.3×10^{-8}	5.7×10^{-8}
dpmethod1	30.6	3.93		0	0

Table 3 Results for data satisfying exactly the pressure equation: deviations from 20°C of the temperatures for the test and reference pressure balances are samples from a Gaussian probability distribution with mean zero and standard deviation of 1°C.

	Mass error (mg)	Mass error (ppm)
1	-204.9111171	6.57674241
2	179.4296806	1.91796423
3	32.1517353	0.29455797
4	31.7706028	0.25466715
5	30.1459108	0.21478330
6	27.2776595	0.17490480
7	23.1658488	0.13503062
8	17.8104786	0.09516007
9	11.2115491	0.05529267
10	3.3690602	0.01542808
11	-5.7169881	0.02443395
12	-16.0465958	0.06429360
13	-27.6197629	0.10415101
14	-40.4364894	0.14400631
15	-54.4967753	0.18385957
16	-69.8006206	0.22371087
17	-86.3480253	0.26356026
18	-104.1389894	0.30340779
19	-123.1735129	0.34325350
20	356.5484042	0.95218720

Table 4 Mass errors for simulating the P - and ΔP -methods

	$A_0(mm^2)$	$\lambda(/MPa) \times 10^6$	$E_m(kg)$	$E_A(ppm)$
wllspe: $w_i = 1$	30.6	3.93	0	0
wllspe: $w_i = 1/p_i$	30.59995898	3.94582223	2.1×10^{-5}	1.2
pmethod0	30.59995890	3.94591294	2.1×10^{-5}	1.2
pmethod1	30.59995896	3.94582896	2.1×10^{-5}	1.2
dpmethod0	30.60013987	3.89359944	7.9×10^{-5}	4.2
dpmethod1	30.60013987	3.89359944	7.9×10^{-5}	4.2

Table 5 Simulating the P - and ΔP -methods: temperatures for the test and reference pressure balances are constant.

	$A_0(mm^2)$	$\lambda(/MPa) \times 10^6$	$E_m(kg)$	$E_A(ppm)$
wllspe: $w_i = 1$	30.6	3.93	0	0
wllspe: $w_i = 1/p_i$	30.59995898	3.94582231	2.1×10^{-5}	1.2
pmethod0	30.59995890	3.94591300	2.1×10^{-5}	1.2
pmethod1	30.59995896	3.94582902	2.1×10^{-5}	1.2
dpmethod0	30.60013987	3.89359889	7.9×10^{-5}	4.2
dpmethod1	30.60013987	3.89359923	7.9×10^{-5}	4.2

Table 6 Simulating the P - and ΔP -methods: deviations from $20^\circ C$ of the temperatures for the test and reference pressure balances are samples from a Gaussian probability distribution with mean zero and standard deviation of $0.1^\circ C$.

i_0	$A_0(mm^2)$	$\lambda(/MPa) \times 10^6$	$E_m(kg)$	$E_A(ppm)$
1	30.60013987	3.89359923	7.9×10^{-5}	4.2
2	30.59991675	3.94081198	9.0×10^{-5}	2.6
3	30.59994619	3.94117320	4.0×10^{-5}	1.6
4	30.59998661	3.93076123	2.0×10^{-5}	0.4
5	30.59999657	3.92826918	1.5×10^{-5}	0.3
6	30.59999937	3.92758045	1.3×10^{-5}	0.3
7	30.59999937	3.92756388	1.3×10^{-5}	0.3
8	30.59999783	3.92788942	1.5×10^{-5}	0.3
9	30.59999519	3.92844822	1.6×10^{-5}	0.3
10	30.59999150	3.92921630	1.9×10^{-5}	0.4
11	30.59998661	3.93021605	2.2×10^{-5}	0.4
12	30.59998020	3.93150957	2.7×10^{-5}	0.6
13	30.59997167	3.93321022	3.4×10^{-5}	0.9
14	30.59995996	3.93551645	4.4×10^{-5}	1.3
15	30.59994316	3.93879191	6.0×10^{-5}	1.8
16	30.59991741	3.94376902	7.9×10^{-5}	2.6
17	30.59987365	3.95216173	1.2×10^{-4}	3.9
18	30.59978470	3.96911848	1.9×10^{-4}	6.6
19	30.59951444	4.02039555	4.2×10^{-4}	15.0
20	30.59921123	4.09300228	6.1×10^{-4}	24.1

Table 7 Results of applying the ΔP -method using different measurements to provide prior knowledge for eliminating c .

	$\text{mean}(E_m)$	$\text{std}(E_m)$	$\text{max}(E_m)$
wllspe	6.9×10^{-6}	3.6×10^{-6}	1.8×10^{-5}
wllspec	8.8×10^{-6}	3.4×10^{-6}	1.8×10^{-5}
pmethod0	1.4×10^{-5}	8.5×10^{-6}	4.3×10^{-5}
dpmethod0	3.1×10^{-6}	2.1×10^{-5}	9.5×10^{-5}

	$\text{mean}(E_A)$	$\text{std}(E_A)$	$\text{max}(E_A)$
wllspe	0.3	0.2	1.1
wllspec	0.9	0.6	3.3
pmethod0	0.7	0.5	2.5
dpmethod0	1.5	1.1	4.8

Table 8 Summary of results for data for which normally distributed random measurement error is added to the values of applied load for the test balance. The unit for the measure E_m of approximation is kg, and the unit for the measure E_A of how well the effective area is estimated is ppm.

	$\text{mean}(E_m)$	$\text{std}(E_m)$	$\text{max}(E_m)$
wllspe	0	0	0
wllspec	5.0×10^{-6}	3.4×10^{-6}	1.8×10^{-5}
pmethod0	1.3×10^{-5}	9.4×10^{-6}	4.6×10^{-5}
dpmethod0	3.3×10^{-5}	2.2×10^{-5}	1.1×10^{-4}

	$\text{mean}(E_A)$	$\text{std}(E_A)$	$\text{max}(E_A)$
wllspe	0	0	0
wllspec	0.9	0.7	3.4
pmethod0	0.8	0.6	2.7
dpmethod0	1.6	1.1	5.2

Table 9 Summary of results for data for which the least-squares estimator returns the true parameter values.

	Mass error (mg)	Mass error (ppm)
1	86.07056466	2.762496058463
2	-111.27542243	1.189451013042
3	-125.29142779	1.147860423035
4	-18.98545246	0.152184060309
5	-150.38065532	1.071432216067
6	64.70472118	0.414889491357
7	195.51114649	1.139612321795
8	63.99616550	0.341927071802
9	291.81482049	1.439160810588
10	-203.11538301	0.930136821629
11	209.83771259	0.896827765400
12	-23.76810230	0.095231684288
13	-149.00927003	0.561896845145
14	-27.83044899	0.099112274478
15	10.12726342	0.034167073497
16	-239.39276584	0.767259344890
17	36.21507201	0.110539452326
18	-13.61323291	0.039661947461
19	-73.70206335	0.205389277804
20	185.77580219	0.496128749358

	Pressure error (Pa)	Pressure error (ppm)
1	-16.78108967	1.67810897
2	21.69869574	0.72328986
3	24.43275967	0.69807885
4	3.70245513	0.09256138
5	29.32770007	0.65172667
6	-12.61937524	0.25238750
7	-38.13210081	0.69331092
8	-12.48220763	0.20803679
9	-56.91964262	0.87568681
10	39.61993660	0.56599909
11	-40.93298523	0.54577314
12	4.63658726	0.05795734
13	29.06943999	0.34199341
14	5.42951756	0.06032797
15	-1.97582786	0.02079819
16	46.70699757	0.46706998
17	-7.06610784	0.06729627
18	2.65625818	0.02414780
19	14.38151459	0.12505665
20	-36.25185611	0.30209880

Table 10 Mass and pressure errors for a data set for which the GDR estimator returns the true parameter values.

	$\text{mean}(E_m)$	$\text{std}(E_m)$	$\text{max}(E_m)$
wllspe	1.9×10^{-5}	2.2×10^{-9}	1.9×10^{-5}
wllspec	2.0×10^{-5}	9.2×10^{-7}	2.6×10^{-5}
pmethod0	2.6×10^{-5}	7.1×10^{-6}	6.3×10^{-5}
dpmethod0	6.9×10^{-5}	4.0×10^{-5}	2.0×10^{-4}

	$\text{mean}(E_A)$	$\text{std}(E_A)$	$\text{max}(E_A)$
wllspe	9.0×10^{-5}	6.7×10^{-5}	3.4×10^{-4}
wllspec	1.7	1.2	6.7
pmethod0	0.9	0.6	3.6
dpmethod0	3.1	2.1	9.3

Table 11 Summary of results for data for which the GDR estimator returns the true parameter values.

	$\text{mean}(E_m)$	$\text{std}(E_m)$	$\text{max}(E_m)$
wllspe	1.5×10^{-4}	2.8×10^{-10}	1.5×10^{-4}
wllspec	1.5×10^{-4}	7.0×10^{-7}	1.5×10^{-4}
pmethod0	1.5×10^{-4}	1.1×10^{-5}	1.9×10^{-4}
dpmethod0	1.6×10^{-4}	4.4×10^{-5}	3.3×10^{-4}

	$\text{mean}(E_A)$	$\text{std}(E_A)$	$\text{max}(E_A)$
wllspe	3.5	3.9×10^{-5}	3.5
wllspec	4.3	0.8	8.9
pmethod0	3.8	0.5	5.3
dpmethod0	4.9	1.6	13.3

Table 12 Summary of results for data with random measurement error and systematic error.

	$\text{mean}(E_m)$	$\text{std}(E_m)$	$\text{max}(E_m)$
wllspe	0	0	0
wllspec	3.9×10^{-6}	2.8×10^{-6}	1.4×10^{-5}
pmethod0	8.6×10^{-6}	6.2×10^{-6}	2.8×10^{-5}
dpmethod0	4.6×10^{-5}	2.9×10^{-5}	1.5×10^{-4}

	$\text{mean}(E_A)$	$\text{std}(E_A)$	$\text{max}(E_A)$
wllspe	0	0	0
wllspec	0.7	0.5	2.4
pmethod0	0.5	0.4	1.8
dpmethod0	3.1	2.1	10.8

Table 13 Summary of results for data defined by a distribution of pressure values that is uniformly spread throughout the pressure range 10 to 120MPa.

	$\text{mean}(E_m)$	$\text{std}(E_m)$	$\text{max}(E_m)$
wllspe	0	0	0
wllspec	5.3×10^{-6}	3.6×10^{-6}	2.0×10^{-5}
pmethod0	1.1×10^{-5}	7.2×10^{-6}	3.5×10^{-5}
dpmethod0	5.9×10^{-5}	3.9×10^{-5}	2.0×10^{-4}

	$\text{mean}(E_A)$	$\text{std}(E_A)$	$\text{max}(E_A)$
wllspe	0	0	0
wllspec	1.1	0.7	3.9
pmethod0	0.7	0.5	2.4
dpmethod0	4.1	2.9	14.2

Table 14 Summary of results for data defined by a distribution of pressure values that is uniformly spread throughout the pressure range 10 to 100MPa.

B SOFTWARE LISTINGS

Disclaimer. The software listed in this appendix has not been fully subjected to NPL's Quality Assurance procedures. No warranty or guarantee applies to this software, and therefore users should satisfy themselves that it meets their requirements.

B.1 LINEAR LEAST-SQUARES METHOD

```
function [A0, lambda, me] = wllspe(p, t, m, w, C, alpha, g)
%
% WLLSPE    Assuming the recorded pressure and temperature values
%           are exact, and there is error in the values of applied
%           load, solve the pressure equation in a weighted least
%           squares sense to determine the effective area parameters
%           for a given pressure balance.
%
% Input parameters:
%   p       Pressure values (MPa).
%   t       Temperature values (degC).
%   m       Applied loads (kg).
%   w       Weighting of values of applied load (wi should be
%           inversely proportional to the standard error for mi).
%   C,
%   alpha   Parameters for pressure balance (see PCONST).
%   g       Gravitational constant (m/s^2).
%
% Output parameters:
%   A0,
%   lambda  Effective area parameters for pressure balance.
%   me      Applied loads determined by the pressure equation
%           and the computed estimates of effective area.
%
% [A0, lambda, me] = wllspe(p, t, m, w, C, alpha, g)
%
% Weighted observation matrix (depending on exact data p and t)
%
%   wt = (1 + alpha.*(t - 20)).*w;
%   A   = [wt.*p   wt.*p.*p];
%
% Weighted right-hand side vector (depending on inexact data m)
%
%   y = w.*(m + C).*g;
%
% Solve for parameters b = [A0   A0*lambda]'
%
%   b = A\y
%
% Extract effective area parameters.
%
%   A0      = b(1);
%   lambda  = b(2)/b(1);
%
% Compute values of applied load that are determined by the
% pressure equation and these estimates.
%
```

```

me = pmass(p, t, A0, lambda, C, alpha, g)
%
%

function [A0, lambda, C, me] = wllspec(p, t, m, w, alpha, g)
%
% WLLSPEC    Assuming the recorded pressure and temperature values
%            are exact, and there is error in the values of applied
%            load, solve the pressure equation in a weighted least
%            squares sense to determine the effective area parameters
%            for a given pressure balance, as well as an estimate of
%            the correction term C to the applied loads for the
%            balance.
%
% Input parameters:
%   p      Pressure values (MPa).
%   t      Temperature values (degC).
%   m      Applied loads (kg).
%   w      Weighting of values of applied load (wi should be
%           inversely proportional to the standard error for mi)
%   alpha  Parameter for pressure balance (see PCONST) used for
%           temperature correction.
%   g      Gravitational constant (m/s^2).
%
% Output parameters:
%   A0,    Effective area parameters for pressure balance.
%   lambda Estimate of correction to applied loads for balance
%   C      Applied loads determined by the pressure equation
%           and the computed estimates of effective area.
%
[A0, lambda, C, me] = wllspec(p, t, m, w, alpha, g)

%
% Weighted observation matrix (depending on exact data p and t)
%
wt = (1 + alpha.*(t - 20)).*w;
A = [wt.*p  wt.*p.*p  -w.*ones(p).*g];
%
% Weighted right-hand side vector (depending on inexact data m)
%
y = w.*m.*g;
%
% Solve for parameters b = [A0  A0*lambda  C]'
%
b = A\y
%
% Extract effective area parameters.
%
A0 = b(1);
lambda = b(2)/b(1);
C = b(3);
%
% Compute values of applied load that are determined by the
% pressure equation and these estimates.
%

```

```

me = pmass(p, t, A0, lambda, C, alpha, g);
%
% -----

```

B.2 P-METHOD

```

function [A0te, lamte, mte] = pmethod0(p, ts, ms, A0s, lams, Cs, alps
                                     tt, mt, Ct, alpt, g)
%
% PMETHOD0 P-method with linearisation for determining the effective
%             area parameters for a given test balance.
%
% Input parameters:
%   p      Pressure values (MPa).
%   ts     Temperature values (degC) for standard balance
%   ms     Applied loads (kg) for standard balance.
%   A0s,   Effective area parameters for standard balance
%   lams,  Cs,
%   alps   Parameters for standard balance (see PCONST).
%   tt     Temperature values (degC) for test balance.
%   mt     Applied loads (kg) for test balance.
%   Ct,   alpt
%   alpt   Parameters for test balance (see PCONST).
%   g      Gravitational constant (m/s^2).
%
% Output parameters:
%   A0te,  Effective area parameters for test balance.
%   lamte  Applied loads determined by the pressure equation
%   mte    and the computed estimates of effective area.
%
% [A0te, lamte, mte] = pmethod0(p, ts, ms, A0s, lams, Cs, alps,
%                               tt, mt, Ct, alpt, g)
%
% Observation matrix (obtained by linearising the function of p).
%
% A = [ones(p)  p];
%
% Right-hand side vector
%
% y1 = (mt + Ct)./(ms + Cs);
% y2 = (1 + alps.*(ts - 20))./(1 + alpt.*(tt - 20))
% y = y1.*y2;
%
% Least squares solution (for parameters derived from linearisation)
%
% b = A\y;
%
% Extract effective area parameters
%
% A0te = A0s*b(1);
% lamte = lams + b(2)/b(1)
%
% Compute values of applied load that are determined by the
% pressure equation and these estimates.

```

```
%
mte = pmass(p, tt, A0te, lamte, Ct, alpt, g);
%
% -----

function [A0te, lamte, mte] = pmethod1(p, ts, ms, A0s, lams, Cs, alps, ...
                                       tt, mt, Ct, alpt, g)

%
% PMETHOD1 P-method for determining the effective area
%            parameters for a given test balance, modified to
%            avoid linearisations.
%
% Input parameters:
%   p      Pressure values (MPa).
%   ts     Temperature values (degC) for standard balance
%   ms     Applied loads (kg) for standard balance.
%   A0s,   Effective area parameters for standard balance.
%   lams   Effective area parameters for standard balance.
%   Cs,    Parameters for standard balance (see PCONST).
%   alps   Parameters for standard balance (see PCONST).
%   tt     Temperature values (degC) for test balance.
%   mt     Applied loads (kg) for test balance.
%   Ct,    Parameters for test balance (see PCONST).
%   alpt   Parameters for test balance (see PCONST).
%   g      Gravitational constant (m/s^2).
%
% Output parameters:
%   A0te,  Effective area parameters for test balance.
%   lamte  Effective area parameters for test balance.
%   mte    Applied loads determined by the pressure equation
%          and the computed estimates of effective area.
%
% [A0te, lamte, mte] = pmethod1(p, ts, ms, A0s, lams, Cs, alps,
%                               tt, mt, Ct, alpt, g)
%
% Observation matrix.
%
% A1 = ones(p)./(1 + lams.*p)
% A2 = p.*A1;
% A = [A1 A2];
%
% Right-hand side vector.
%
% y1 = (mt + Ct)./(ms + Cs);
% y2 = (1 + alps.*(ts - 20))./(1 + alpt.*(tt - 20));
% y = y1.*y2;
%
% Least squares solution b = [A0t/A0s lamt*A0t/A0s]'
%
% b = A\y;
%
% Extract effective area parameters.
%
% A0te = A0s*b(1);
% lamte = A0s*b(2)/A0te;
```

```
%
% Compute values of applied load that are determined by the
% pressure equation and these estimates.
%
mte = pmass(p, tt, A0te, lamte, Ct, alpt, g);
%
%
```

B.3 ΔP -METHOD

```
function [A0te, lamte, mte] = dpmethod0(p, tt, mt, Ct, alpt, g)
%
% DPMETHOD0 DeltaP-method for determining the effective area
% parameters for a given test balance.
%
% Input parameters:
%   p      Pressure values (MPa).
%   tt     Temperature values (degC) for test balance.
%   mt     Applied loads (kg) for test balance.
%   Ct,
%   alpt   Parameter for test balance (see PCONST).
%   g      Gravitational constant (m/s^2).
%
% Output parameters:
%   A0te,
%   lamte  Effective area parameters for test balance.
%   mte    Applied loads determined by the pressure equation
%           and the computed estimates of effective area.
%
[A0te, lamte, mte] = dpmethod0(p, tt, mt, Ct, alpt, g)

%
% Number of measurements
%
mp = length(p);
%
% Observation matrix.
%
ps = p(1) + p(2:mp);
A = [ones(ps) ps];
%
% Right-hand side vector
%
dm = mt(2:mp) - mt(1);
dp = p(2:mp) - p(1);
y1 = dm.*g./(1 + alpt.*(tt(2:mp) - 20));
y2 = dp - alpt.*(tt(1) - tt(2:mp)).*p(1);
y = y1./y2;
%
% Least squares solution
%
b = A\y;
%
% Extract effective area parameters
%
A0te = b(1);
lamte = b(2)/b(1);
```

```

%
% Compute values of applied load that are determined by the
% pressure equation and these estimates.
%
mte = pmass(p, tt, A0te, lamte, Ct, alpt, g)
%
%

function [A0te, lamte, mte] = dpmethod1(p, tt, mt, Ct, alpt, g)
%
% DPMETHOD1 Simulation of the DeltaP-method for determining
% the effective area parameters for a given test balance
% by solving a weighted least squares problem derived
% from the pressure equation. The correction term Ct
% is implicitly eliminated by using information derived
% from the first measurement.
%
% Input parameters:
% p Pressure values (MPa).
% tt Temperature values (degC) for test balance.
% mt Applied loads (kg) for test balance.
% Ct,
% alpt Parameters for test balance (see PCONST).
% g Gravitational constant (m/s^2).
%
% Output parameters:
% A0te,
% lamte Effective area parameters for test balance.
% mte Applied loads determined by the pressure equation
% and the computed estimates of effective area.
%
[A0te, lamte, mte] = dpmethod1(p, tt, mt, Ct, alpt, g)

%
% Number of measurements.
%
mp = length(p)

%
% Weighted observation matrix
%
dp = p(2:mp) - p(1);
sp = p(2:mp) + p(1);
w = ones(dp)./dp;
A = [w.*dp w.*dp.*sp];

%
% Weighted right-hand side vector
%
y1 = (mt(2:mp) + Ct)./(1 + alpt.*(tt(2:mp) - 20));
y2 = (mt(1) + Ct)./(1 + alpt.*(tt(1) - 20));
y = w.*(y1 - y2).*g;

%
% Least squares solution.
%
b = A\y;

%
% Extract effective area parameters.

```

```

%
%   A0te = b(1);
%   lamte = b(2)/b(1);
%
%   Compute values of applied load that are determined by the
%   pressure equation and these estimates.
%
%   mte = pmass(p, tt, A0te, lamte, Ct, alpt, g);
%
%

function [A0te, lamte, mte] = dpmethodi(i0, p, tt, mt, Ct, alpt, g)
%
%   DPMETHODI Simulation of the DeltaP-method for determining
%               the effective area parameters for a given test balance
%               by solving a weighted least squares problem derived
%               from the pressure equation. In this version, the
%               correction term Ct is implicitly removed using
%               information from the (i0)th measurement.
%
%   Input parameters:
%       i0      Index of measurement used to eliminate Ct.
%       p       Pressure values (MPa).
%       tt      Temperature values (degC) for test balance.
%       mt      Applied loads (kg) for test balance.
%       Ct,
%       alpt    Parameters for test balance (see PCONST).
%       g       Gravitational constant (m/s^2).
%
%   Output parameters:
%       A0te,
%       lamte   Effective area parameters for test balance.
%       mte     Applied loads determined by the pressure equation
%               and the computed estimates of effective area.
%
%   [A0te, lamte, mte] = dpmethodi(i0, p, tt, mt, Ct, alpt, g)

%
%   Number of measurements
%
%   mp = length(p);
%
%   Remove (i0)th measurement from data.
%
%   pb = p; pb(i0) = [];
%   tb = tt; tb(i0) = [];
%   mb = mt; mb(i0) = [];
%
%   Weighted observation matrix
%
%   dp = pb - p(i0);
%   sp = pb + p(i0);
%   w = ones(dp)./dp;
%   A = [w.*dp    w.*dp.*sp];
%
%   Weighted right-hand side vector

```

```
%
y1 = (mb      + Ct)./(1 + alpt.*(tb      - 20));
y2 = (mt(i0) + Ct)./(1 + alpt.*(tt(i0) - 20));
y  = w.*(y1 - y2).*g;
%
% Least squares solution
%
b = A\y;
%
% Extract effective area parameters.
%
A0te = b(1);
lamte = b(2)/b(1)
%
% Compute values of applied load that are determined by the
% pressure equation and these estimates.
%
mte = pmass(p, tt, A0te, lamte, Ct, alpt, g);
%
%
```

B.4 GENERATION OF EXACT DATA

```
function [A0s, lams, Cs, alps, A0t, lamt, Ct, alpt, g] = pconst
%
% PCONST  Matlab function for specifying constants for a standard
%          and a test pressure balance.
%
%
% Output parameters:
%   A0s, lams, Cs, alps  Parameters for standard balance.
%   A0t, lamt, Ct, alpt  Parameters for test balance.
%   rhoa, g              Air density and gravitational constant
%
% [A0s, lams, Cs, alps, A0t, lamt, Ct, alpt, g] = pconst
%
% Gravitational constant (m/s^2)
%
g      = 9.8100;
%
% Data for standard pressure balance:
%   effective area parameters (A0s mm^2 and lams /MPa),
%   correction to applied load (Cs kg),
%   temperature coefficient (alps /degC),
%
A0s = 3.07e+01;
lams = 4.10e-06;
Cs = 3.67e-02;
alps = 2.34e-05;
%
% Data for test balance
%
A0t = 3.06e+01;
lamt = 3.93e-06;
Ct = 3.70e-02;
alpt = 2.34e-05;
```



```

%
% -----
function m = pmass(p, t, A0, lambda, C, alpha, g)
%
% PMASS      Given a vector of pressures, a vector of temperatures
%             and the parameters for a pressure balance, PMASS returns
%             the vector of applied loads for which the pressure
%             equation is satisfied with this data.
%
% Input parameters:
%   p        Pressure values (MPa).
%   t        Temperature values (degC).
%   A0,
%   lambda,
%   C,
%   alpha    Parameters for pressure balance (see PCONST).
%   g        Gravitational constant (m/s^2).
%
% Output parameters:
%   m        Applied loads (kg) including air density correction.
%
% m = pmass(p, t, A0, lambda, C, alpha, g);
%
% Applied loads are computed from a rearrangement of the
% pressure equation.
%
% m = (A0/g)*(p.*(1 + lambda*p).*(1 + alpha*(t - 20))) - C;
%
% -----

```

B.5 GDR DATA GENERATION

```

function [x, y, t, f, fd, A] = ndgsgdr(xs, a, alpha, beta, nd, ...
                                       fgeval, p1, p2, p3, p4, p5)
%
% NDGSGDR      Null space data generator for simple generalised
%              distance regression:
%
%              min_{xs, a} sum_i { alpha_i^2 (x_i - xs_i)^2
%                                + beta_i^2 (y_i - ys_i)^2 },
%
%              where ys_i = f(xs_i, a).
%
% Input
%   xs - m x 1 array of x-ordinates.
%   a  - n x 1 array of parameters for f.
%   alpha, beta - m x 1 arrays of weights.
%   nd  - 1 x 1 array specifying the norms of 1 perturbations.
%   fgeval - string name of a function and gradient evaluation
%            module specified by
%
%           [ f, fd, A ] = fgeval(xs, a, p1, p2, ...).
%
%

```

```

%   p1, p2,... -      parameters to be passed to fgeval.
%
% Output
%       x, y - m x 1 arrays of x and y coordinates. {(x(:, k), y(:,k))}
%               is a set of data points for which the GDR solution
%               is a with GDR norm nd(k).
%       t - m x 1 array of perturbations.
%       f - m x 1 array of f evaluated at xs.
%       fd - m x 1 array of derivatives df/dx evaluated at (xs_i, a)
%       A - m x n Jacobian matrix df/da_j evaluated at (xs_i, a).
%
% [x, y, t, f, fd, A] = ndgsgdr(xs, a, alpha, beta, nd, ...
%                               'fgeval', p1, p2, p3, p4, p5)

```

```

function [ f, fd, A ] = pbfgeval(ps, a, phi, c, g)
%
% PBFGEVAL      Pressure balance function and gradient evaluation
%               for a quadratic model:
%
%                $f = (1 + \phi) \cdot ps \cdot (a(1) + a(2) \cdot ps) / g - c.$ 
%
% Input
%       ps - m x 1 array of pressures.
%       a - 2 x 1 array of coefficients.
%       phi - m x 1 array of temperature correction terms.
%       c - 1 x 1 mass correction term.
%       g - 1 x 1 gravitational constant.
%
% Output
%       f - m x 1 array of mass values.
%       fd - m x 1 array of derivatives df/dp evaluated at ps_i.
%       A - m x 1 array of derivatives df/da_j evaluated at ps_i
%
% [ f, fd, A ] = pbfgeval(ps, a, phi, c, g)
%
% Function values.
%
%  $f = (1 + \phi) \cdot ps \cdot (a(1) + a(2) \cdot ps) / g - c;$ 
%
% Derivatives with respect to pressure values
%
%  $fd = (1 + \phi) \cdot (a(1) + 2 \cdot a(2) \cdot ps) / g;$ 
%
% Derivatives with respect to effective area parameters
%
%  $h = (1 + \phi) \cdot ps / g;$ 
%  $A = [h \ h \cdot ps];$ 
%
% -----

```