Final Report on Key Comparison CCM.P-K4 in Absolute Pressure from 1 Pa to 1000 Pa

A.P. Miiller[1], M. Bergoglio[2], N. Bignell[3], K.M.K. Fen[3], S. S. Hong[4], K. Jousten[5], P. Mohan[6], F. J. Redgrave[7], and M. Sardi[2]

January 22, 2002

ABSTRACT

This report describes a CCM key comparison of low absolute-pressure standards at seven National Metrology Institutes (NMIs) that was carried out during the period March 1998 to September 1999 in order to determine their degrees of equivalence at pressures in the range 1 Pa to 1000 Pa. The primary standards, which represent two principal measurement methods, included five liquid-column manometers and four static expansion systems. The transfer standard package consisted of four high-precision pressure transducers, two capacitance diaphragm gauges to provide high resolution at low pressures, and two resonant silicon gauges to provide the required calibration stability. Two nominally identical transfer packages were used to reduce the time required for the measurements, with Package A being circulated among laboratories in the European region (IMGC, NPL-UK, and PTB) and Package B being circulated among laboratories in the Asia-Pacific region (CSIRO, KRISS, and NPL-I). The results obtained with different transfer packages were normalized by using data obtained from simultaneous calibrations of the two packages at the pilot laboratory (NIST). The degrees of equivalence of the measurement standards were determined in two ways, deviations from key comparison reference values and pairwise differences between these deviations. Apart from results from one NMI that were identified as outliers, the absolute pressure standards of the seven participating NMIs were generally found to be equivalent and the results revealed no significant relative bias between the two principal methods tested by this comparison.

[1] NIST: National Institute of Standards and Technology, Gaithersburg, MD 20899 USA
[2] IMGC-CNR: Istituto di Metrologia "G. Colonnetti" of CNR, Torino, 10135 Italy
[3] CSIRO: Commonwealth Scientific and Industrial Research Organization, Lindfield, N.S.W. 2070 Australia
[4] KRISS: Korea Research Institute of Standards and Science, Taejon, Republic of Korea
[5] PTB: Physikalisch Technische Bundesanstalt, Berlin, Germany
[6] NPL-I: National Physical Laboratory, New Delhi-110012 India
[7] NPL-UK: National Physical Laboratory, Teddington, Middlesex United Kingdom TW110LW

CONTENTS

1. INTRODUCTION

In May 1996 the Comité Consultatif pour la Masse et les grandeurs apparentéès[8] (CCM) approved proposals by the pressure working groups that identified six comparisons in pressure, the relevant ranges, the transfer standards to be used, and the pilot laboratories. The objective of these comparisons, which were seen as a scientific exercise, was to ascertain the relative performance of primary pressure standards developed at selected National Metrology Institutes (NMIs). However, with the signing of the Mutual Recognition Arrangement (MRA) [1] by NMIs of Member States of the Metre Convention in October 1999, it was agreed that the six comparisons would serve as *key comparisons* as provided for in the MRA. A major objective of the MRA is to establish the degree of equivalence of national measurement standards maintained by NMIs through key comparisons that test principal measurement methods in the field.

One of the six key comparisons was in absolute pressure covering the range 1 Pa to 1000 Pa, which it was agreed would be piloted by the National Institute of Standards and Technology (NIST) using high-precision pressure transducers as transfer standards. The participants in the comparison were given the option of extending the range down to 0.1 Pa and up to 10,000 Pa, although extensions to the range would not necessarily be included in the BIPM (Bureau International des Poids et Mesures) database.

During years leading up to the start of the comparison, different types of pressure transducers were tested and evaluated at the pilot laboratory to find a transducer with sufficient pressure resolution, long-term calibration stability, and resistance to over-pressure and mechanical shock that could be used as the transfer standard. It was found that no one type of transducer could satisfy all requirements but rather a combination of two types, capacitance diaphragm gauges to provide high resolution and resonant silicon gauges to provide calibration stability.

This report summarizes the calibrations of the transfer standards carried out at seven NMIs during the period March 1998 and September 1999. Two nominally identical transfer standard packages were used in this comparison to reduce the time required to complete all measurements, with Package A being circulated among laboratories in the European region (IMGC, NPL-UK, and PTB) and Package B being circulated among laboratories in the Asia-Pacific region (CSIRO, KRISS, and NPL-I). Results from the two regions were normalized by using data obtained during simultaneous calibrations of the two packages at the pilot laboratory.

The following sections provide brief descriptions of the primary standards, the design and construction of the transfer standard packages, the organization and chronology of the comparison, and the general calibration procedure required by the protocol. Methods for reduction and analysis of the calibration data were chosen to provide, as much as possible, uniform treatment of the results from individual laboratories, whether they were the pilot or another participant laboratory.

2. PRIMARY STANDARDS

The principal measurement methods tested by this comparison involved two types of primary standards: static expansion systems, which are pressure generators; and liquid-column manometers, which are pressure measurers. Four participants (IMGC, NPL-I, NPL-UK, and PTB) used static expansion systems as their primary standards and four participants used different types of manometers in which liquid-column heights were measured either by laser interferometry (CSIRO and IMGC) or by ultrasonic interferometry (KRISS and NIST).

2.1. STATIC EXPANSION SYSTEM AT THE IMGC

The static expansion system at the IMGC is a modification of that described in reference [2], the principal difference being the addition of a third volume as described in reference [3]. The system consists of three volumes, nominally 10 mL, 500 mL and 50 L, the largest volume being the calibration chamber. The different expansion ratios are measured, and are periodically determined, by application of the multiple-expansion method. The initial pressures between 1 kPa and 100 kPa are measured by secondary transfer standards directly traceable to the HG5 mercury manometer (see Section 2.6). The base pressure, which is

[8] Consultative Committee for Mass and Related Quantities.

obtained by a turbo pump, is in the range of 10^{-6} Pa. The relative combined uncertainty of the system for the pressure range 0.1 Pa to100 Pa is 2.1 x 10^{-3} Pa when volumes added to the system by gauges to be calibrated can be disregarded. In the present comparison, the additional volume of the transfer standard gauges and associated plumbing could not be disregarded and so the added volume was measured using a spinning rotor gauge. This procedure increased the relative combined uncertainty to 3 x 10^{-3}.

2.2. STATIC EXPANSION SYSTEM AT THE NPL-I

The primary standard at the NPL-I used in the key comparison is a static expansion system in which there are two initial volumes v_1, v_2 (nominally 25 mL and 384 mL, respectively) and a large chamber V_L with a nominal volume of 72 L [4 - 7]. V_L can be evacuated to base pressures of 10^{-7} Pa using a diffusion pump and a liquid nitrogen trap equipped with an isolation valve. The initial pressures for the comparison were measured by means of a 110-kPa quartz-spiral Bourdon gauge calibrated against an ultrasonic interferometer manometer. For generating the pressure points 1 Pa, 3 Pa, 10 Pa, and 30 Pa the expansion scheme v_1 to ($v_1 + V_L$) was adopted with initial pressures ranging from 2800 Pa to 85000 Pa. The pressure points 100 Pa and 300 Pa were generated by using the expansion scheme v_2 to ($v_2 + V_L$) with initial pressures of 18900 and 57000 Pa, respectively. The final pressure point of 1000 Pa was generated by using the successive expansion method with two expansions from v_2 to ($v_2 + V_L$). Platinum resistance thermometers inserted into the different chambers were used to measure the temperature of the gas before and after expansion. The volume ratios of the different chambers have been determined both by gravimetry (filling the different chambers with triple distilled water) and also by the method of successive expansion [4, 8].

2.3. STATIC EXPANSION SYSTEM AT THE NPL-UK

The medium vacuum standard (SEA III) at the NPL-UK is a three-stage non-bakeable static expansion system with a 50-L calibration chamber. By varying the initial pressure and the number of stages of expansion, calculable pressures between 1.5×10^{-2} Pa and 2×10^{3} Pa may be generated. There is a choice of two small vessels from which gas may be expanded into the calibration chamber and this enables a greater range of pressures to be generated from a given range of initial pressures. The pressure of the initial gas sample is measured using a quartz Bourdon tube gauge. The pressure generated is calculated from knowledge of the initial pressure, the ratio of the volumes and the gas temperatures. The ratios of the volumes are determined using Elliott's [9] experimental procedure of repeated expansions and are calculated using the iterative method described by Redgrave et al [8].

2.4. STATIC EXPANSION SYSTEM AT THE PTB

The primary standard of the PTB is a static expansion system, called SE2, in which pressures are generated by expanding gas of known pressure from two alternative small volumes of nominally 0.1 L and 1 L directly into a volume of 100 L. It is also possible to carry out two expansions in series with intermediate nominal volumes of 100 L and 1 L. The regular operational range of SE2 is 0.1 Pa up to 1 kPa. The system is described in detail in references [10 - 13].

2.5. LASER INTERFEROMETER MANOMETER AT THE CSIRO

The manometer consists of a mercury U-tube in which the surfaces are the reflectors of a Michelson interferometer [14]. To determine the surface position, tungsten-weighted cat's eye floats are used as reflectors for the laser light. Sloping sides in the float are used to produce a flat mercury surface.

2.6. HG5 LASER INTERFEROMETER MANOMETER AT THE IMGC

The HG5 mercury manometer is the primary pressure standard of the IMGC in the barometric range up to 120 kPa and it can operate in both absolute and differential modes. A full description of HG5 and the discussion of the uncertainties are given in reference [15]. The HG5 is a laser interferometer manometer that essentially consists of two interconnected 60-mm bore, 1-m long glass tubes forming the U-tube, which is placed in a temperature-controlled water bath. The mercury temperature is measured by two

platinum resistance thermometers (PRTs) installed coaxially at the base of the columns. The measurement of the vertical displacements of the mercury menisci is made with a single-beam interferometer. Corner cube reflectors mounted on very lightweight floats, thin glass disks that float on both mercury menisci, reflect the two vertical laser beams. Increased accuracy at pressures up to 13 kPa is obtained by focusing and directly reflecting the laser beams from the mercury menisci using lenses mounted on the floats in a cat's-eye arrangement. Such floats were used for all measurements during the present comparison.

2.7. ULTRASONIC INTERFEROMETER MANOMETER AT THE KRISS

The primary standard at the KRISS for this key comparison is a mercury ultrasonic interferometer manometer that was assembled and evaluated as an international cooperation project between the KRISS and the NIST beginning in 1988. The manometer [16] has an operating range of 10 Pa to 133 kPa and its design and operation are based on the mercury ultrasonic manometers developed at the NIST [17, 18], which are described in the next section.

2.8. ULTRASONIC INTERFEROMETER MANOMETERS AT THE NIST

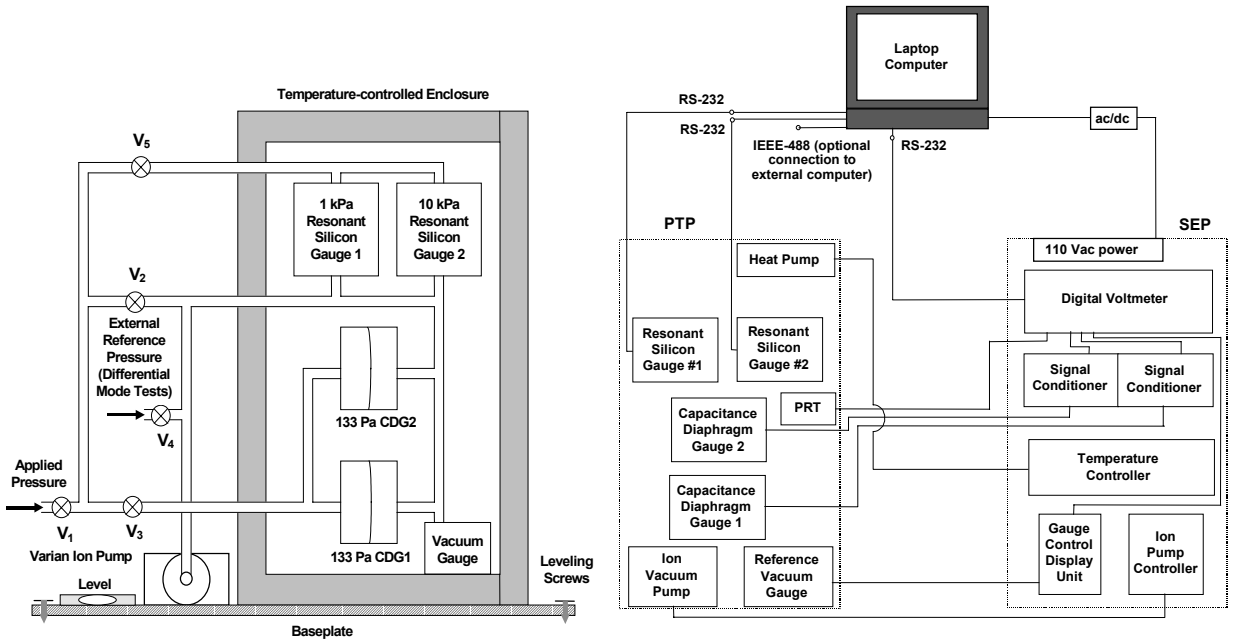
The primary standards at the NIST used in this key comparison are two Ultrasonic Interferometer Manometers (UIMs), a mercury UIM with a full-scale range of 160 kPa [17, 18] and an oil UIM [19] with a full-scale range of 140 Pa. The unique feature of these manometers is that changes in height of the liquid columns are determined by an ultrasonic technique. A transducer at the bottom of each liquid column generates a pulse of ultrasound that propagates vertically up the column, is reflected from the liquid-gas interface, and returns to be detected by the transducer. The length of the column, which is proportional to the change in phase of the returned signal, is determined from the phase change and the velocity of the ultrasound [20]. The manometers have large-diameter (75 mm – Hg UIM; 100 mm – oil UIM) liquid surfaces to minimize capillary effects, thermal shields to stabilize the temperature and minimize its gradients, and high-vacuum techniques to minimize leaks and pressure gradients. The mercury UIM employs a "W" or three-column design to correct for possible tilt. The oil UIM uses a four-column design equivalent to two parallel manometers that also function as orthogonal tilt meters.

3. TRANSFER STANDARDS

On the basis of earlier comprehensive reviews of pressure transducer performance [21, 22], two types of gauges were selected as the transfer standards, namely, resonant silicon gauges (RSGs) for their good long-term stability and capacitance diaphragm gauges (CDGs) for their high-resolution. The RSGs are a new type of MEMS (MicroElectroMechanical Systems) sensor that have excellent calibration stability, are resistant to mechanical shock, and are only moderately susceptible to overpressure although they are rather sensitive to tilt (~ 0.4 Pa/mrad). However they lack sufficient pressure resolution to cover the entire range of the comparison. The CDGs have superior pressure resolution and, because of their all-metal construction, are rugged and resistant to over-pressure but lack the desired calibration stability required by the comparison. The solution was to develop a transfer standard package using both types of gauges, two CDGs to provide redundancy and high resolution at low pressures, and two RSGs to provide redundancy and excellent calibration stability. Good calibration stability was accomplished over the entire pressure range by re-scaling the CDG response to that of the RSGs at an overlapping pressure.

The two RSGs selected for the comparison had full-scale ranges of 1000 Pa and 10,000 Pa and were combined with two CDGs each with a full-scale range of 133 Pa. Since the RSGs were available only as differential units, the decision was made to use differential CDGs as well and to use an ion vacuum pump to provide the near-vacuum reference pressure required for a comparison in absolute pressure. This configuration had the added flexibility that the same transfer standard could also be used for another key comparison in differential pressure (CCM.P-K5) covering the same pressure range.

The transfer standard package consisted of three parts, a pressure transducer package (PTP), a support electronics package (SEP), and a laptop computer (see Figure 1). The PTP included four differential transducers housed in a temperature-controlled enclosure, a calibrated 100-ohm platinum resistance thermometer (PRT) to measure the interior temperature of the enclosure, and an ion vacuum pump and reference-pressure vacuum gauge for the absolute mode calibrations. All-metal plumbing was



(a) (b)

Figure 1. Schematic diagram of (a) the pressure transducer package (PTP) and (b) the electrical connections between the PTP, the support electronics package (SEP) and the laptop computer.

used throughout the PTP including five metal bellows-sealed valves and metal bellows connections to each transducer to minimize mechanical strain. The valves included external isolation valves V_1 and V_4, internal isolation valves for the CDGs (V_3) and RSGs (V_5), and an internal bypass valve V_2 between the pressure and reference side of the gauges. The gauges and internal plumbing were maintained under vacuum during shipment or storage, but with all internal valves open to avoid overpressurization of the gauges in the event of a leak to atmosphere. The tilt orientation of the PTP during calibration of the RSGs was monitored by means of sensitive bubble levels mounted on the PTP base plate and any observed changes were corrected using the leveling screws.

The SEP included a temperature controller for the transducer enclosure, signal-conditioning electronics for the CDGs, a controller for the ion vacuum pump, and a digital voltmeter (DVM) for digitizing analog signals from the CDGs, the calibrated PRT, and the reference vacuum gauge. The enclosure temperature was controlled by means of a heat pump and a Wheatstone bridge mounted inside the enclosure where the bridge included an uncalibrated PRT and an adjustable resistor in two of its arms (not shown in Figure 1b).

A laptop computer was used for controlling the acquisition of data from the RSGs and the DVM during calibration. The time required to obtain one set of readings was approximately 55 seconds. Because of their accuracy (~ 1 part in 10^4), the readings of the RSGs were multiplied by a scale factor before display and storage on the laptop computer in order to increase the level of confidentiality for the pilot laboratory data. The RSG data submitted by the participants were multiplied by 1/(scale factor) during subsequent data reduction in order to restore the original readings.

For interlaboratory shipment, the PTP and SEP (with the laptop) were mounted in commercial containers that were specially designed for vibration and shock isolation.

4. ORGANIZATION OF THE KEY COMPARISON

The present key comparison in absolute pressure (CCM.P-K4) was organized in conjunction with another key comparison in differential pressure (CCM.P-K5) covering the same range in order to minimize the time required for completion of both comparisons. Two nominally identical transfer standard packages were developed for use in either absolute or differential mode. Transfer standard package A was circulated through the European region while transfer standard package B was circulated through the Asian-Pacific region with repeat calibrations at the pilot laboratory (NIST) in an approximate "star" sequence.

4.1. CHRONOLOGY OF THE MEASUREMENTS

Table 1 presents the actual chronology of calibrations during the measurement phase of both comparisons. The start and end dates refer to the measurement time period during which calibration data was taken at each NMI. The total time required to complete the measurement phase of the absolute pressure comparison was eighteen months, which is approximately three months longer than originally projected due to unanticipated problems (see next section).

Table 1. Chronology of measurements during two key comparisons, one in absolute pressure (in bold) the other in differential pressure.

NMI	Transfer Std Package	Calibration Start Date	Calibration End Date	Calibration Mode
		European Region		
NIST #1	**A**	**March 17, 1998**	**April 3, 1998**	**Absolute**
PTB	**A**	**May 28, 1998**	**June 4, 1998**	**Absolute**
NIST #2	**A**	**July 8, 1998**	**July 24, 1998**	**Absolute**
		July 30, 1998	Aug 5, 1998	Differential
NPL-UK	**A**	Oct 29, 1998	Nov 4, 1998	Differential
		Nov 17, 1998	**Nov 25, 1998**	**Absolute**
IMGC	**A**	**Jan 19, 1999**	**Feb 8, 1999**	**Absolute**
		Feb 18, 1999	Feb 23, 1999	Differential
NIST #3	**A**	**April 23, 1999**	**May 6, 1999**	**Absolute**
		May 11, 1999	May 17,1999	Differential
		Asia-Pacific Region		
NIST #2	**B**	**July 8, 1998**	**July 24, 1998**	**Absolute**
		July 30, 1998	Aug 5, 1998	Differential
MSL-NZ†	B	Oct 30, 1998	Nov 5, 1998	Differential
NPL-I	**B**	**Jan 1, 1999**	**Jan 14, 1999**	**Absolute**
CSIRO	**B**	**Feb 24, 1999**	**March 3, 1999**	**Absolute**
		March 11, 1999	March 18, 1999	Differential
NIST #3	**B**	**April 23, 1999**	**May 6, 1999**	**Absolute**
		May 11, 1999	May 17,1999	Differential
KRISS	**B**	**June 15, 1999**	**June 22, 1999**	**Absolute**
NIST #4	**B**	**Aug 23, 1999**	**Sept 10, 1999**	**Absolute**

† Measurement Standards Laboratory- New Zealand

The sequence of calibrations for the **key comparison in absolute pressure** included simultaneous calibrations of the two transfer standard packages during the second and third calibration cycles at the pilot laboratory (NIST #2 and NIST #3) as follows:

Package A: **NIST #1** => PTB => **NIST #2** => NPL-UK => IMGC => **NIST #3**

Package B: **NIST #2** => NPL-I => CSIRO => **NIST #3** => KRISS => **NIST #4**

4.2. PROBLEMS DURING THE COMPARISONS

There were several problems during the course of the comparisons, ranging from instrument malfunction during calibration to severe damage to the transfer standard package during shipment between NMIs. In most cases instrument failures were remedied by replacement with an equivalent unit. However a potentially more serious instrument failure was the rare but intermittent malfunction of one of the 10 kPa RSGs, first observed during the initial evaluation of these gauges. The manufacturer was unable to diagnose the problem but did send a new processor board as a backup. At the start of the first calibration of Package B (NIST #2) the offending RSG began to malfunction again and the processor board was replaced. Unexpectedly during calibrations at the IMGC, the 10 kPa RSG in Package A also began to exhibit the same behavior after completing only one run of the absolute mode calibrations but began operating normally again when used for the differential mode calibrations. The gauge continued to operate normally after Package A was shipped to the pilot laboratory for its third and final calibration (NIST #3).

Very rough handling of the transfer standard containers during shipment caused the most severe problems and contributed to significant delays in completing the comparisons. When Package B arrived at the NPL-I, the gauges in the PTP were found to be at atmosphere due to the rupture of a metal bellows inside the thermal enclosure. A replacement bellows was fabricated at the NIST and sent to the NPL-I to enable the repair. The most serious damage to Package B however occurred during shipment from the CSIRO to the pilot laboratory and was consistent with penetration of the PTP container by a forklift truck. The force of the impact was sufficient to dismount both CDGs and rupture several metal bellows inside the thermal enclosure. Remarkably after Package B was repaired and re-calibrated (NIST #3), the gauges did not show any unusual shifts in their calibration (see Section 6.3).

5. GENERAL CALIBRATION PROCEDURE

The general procedure for the key comparison required that each laboratory calibrate the transfer standard with nitrogen gas at the following nominal absolute pressures in ascending order: 1 Pa, 3 Pa, 10 Pa, 30 Pa, 100 Pa, 300 Pa, and 1000 Pa. The nominal reference or base pressure provided by the ion vacuum pump was to be $\sim 10^{-3}$ Pa. The actual absolute pressures realized at the transfer standard gauges by the participant's pressure standard were to be within 2 parts in 100 of the target pressures. Optional calibration data[9] could also be taken at 0.1 Pa, 0.3 Pa, and at 3000 Pa, 10,000 Pa.

A total of **five calibration runs** were required, with each run taken on a different day. Within a calibration run, **five repeat sets** of pressure and temperature readings of the transfer standard and primary standard were required at each target pressure. At the beginning of each calibration run, **ten repeat sets** of zero-pressure readings for the transfer standard gauges were required to be taken with the PTP isolated from the participant's calibration system (valves V_1 and V_4 closed) and with internal isolation valves V_3 and V_5 and bypass valve V_2 open. These data were needed to correct calibration data obtained with liquid-column manometers for zero-pressure offsets. An additional ten repeat sets of zero-pressure readings were to be taken at the end of each run in order to monitor zero drift in the four transducers during calibration. The calibration procedure also included the option of recording **five sets** of zero-offset readings for the gauges just prior to establishing each target pressure. These readings, which were taken with the external and internal isolation valves of PTP open and bypass valve V_2 closed, were needed to correct zero offsets in calibration data obtained with static expansion systems.

The format for reporting calibration data followed the measurement sequence dictated by the data acquisition software. The sequence for each set of associated readings of the transfer standard and the participant's primary standard was:

$$\textit{Set No.} \quad p_{CDG1} \quad p_{CDG2} \quad p_{RSG1} \quad p_{RSG2} \quad p_{REF} \quad t_{PRT} \quad P_{STD} \quad t_{STD}$$

where the meaning of subscripts for pressures p (gauges), P (primary standard) and temperatures t are self-explanatory. All calibration data were transmitted to the pilot laboratory in the form of spreadsheet files, which greatly facilitated the analyses of a rather voluminous amount of data.

[9] Optional data, which were taken by NIST, NPL-UK, and PTB at pressures below 1 Pa, and by CSIRO, KRISS, and NIST at pressures above 1000 Pa, are not included in this report.

6. REDUCTION AND ANALYSIS OF THE REPORTED DATA

The reduction and analysis of the key comparison data required that several factors be addressed. These included zero-pressure offsets (Section 6.1), thermal transpiration effects (Section 6.2), deviations of the actual pressures realized from the target pressures (Section 6.3), relatively large calibration shifts in the capacitance diaphragm gauges (Section 6.4), and normalization of data from two different transfer standard packages (Section 6.5). Methods for estimating uncertainties (Sections 6.6 and 6.7) and for evaluating degrees of equivalence (Section 6.8) are also described.

6.1. CORRECTIONS FOR ZERO-PRESSURE OFFSETS

The first step in reducing the comparison data was to correct the readings of each gauge *i* for its zero-pressure offset. The index *i* is equal to either 1 or 2 and refers to either, CDG1 and CDG2, or RSG1 and RSG2 (see Figure 1). At a given target pressure during calibration run *k,* the corrected reading of gauge *i* for repeat set *l* is given by:

$$p_{ikl} = p_{Gikl} - \left\langle p_{Gik0} \right\rangle_{10} + p_{REFk\,l} \qquad \text{for liquid-column manometer data} \qquad (1a)$$

and

$$p_{ikl} = p_{Gikl} - \left\langle p_{Gik0} \right\rangle_{5} \qquad \text{for static expansion system data} \qquad (1b)$$

where p_{Gikl} is the uncorrected gauge reading, $\left\langle p_{Gik0} \right\rangle_{10}$ is the mean of 10 zero-pressure readings taken just prior to the start of calibration run *k*, p_{REFkl} is the reference pressure reading during repeat set *l*, and $\left\langle p_{Gik0} \right\rangle_{5}$ is the mean of 5 zero-offset readings taken just prior to realizing each target pressure.

6.2. CORRECTIONS FOR THERMAL TRANSPIRATION EFFECTS

The difference in temperature of a primary standard and the transfer standard gauges can give rise to significant thermal transpiration effects at low pressures [23]. In the present comparison the magnitude of this effect will vary since the primary standards were operated at somewhat different temperatures (see Figure 2). The effect of different operating temperatures was minimized by determining the pressure that a primary standard would measure/generate if it were operating at the same temperature as the transfer standard gauges.

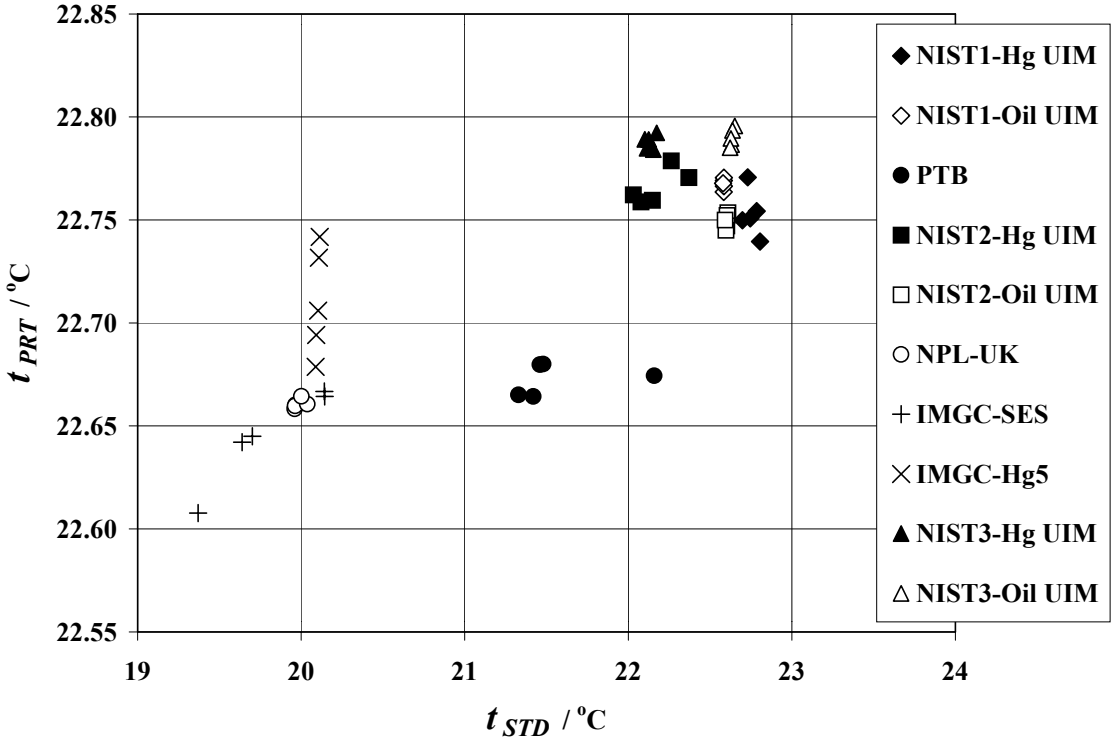
At a given target pressure during calibration run *k*, the corrected reading of primary standard *j* for repeat set *l* is given to a good approximation by the Takaishi-Sensui equation [24]:

$$P_{jkl} = P_{Sjkl} f_{tt} = P_{Sjkl} \times \left(aY^2 + bY + cY^{1/2} + 1\right) \Big/ \left(aY^2 + bY + cY^{1/2} + (T_{Sjkl}/T_{Gikl})^{1/2}\right) \qquad (2a)$$

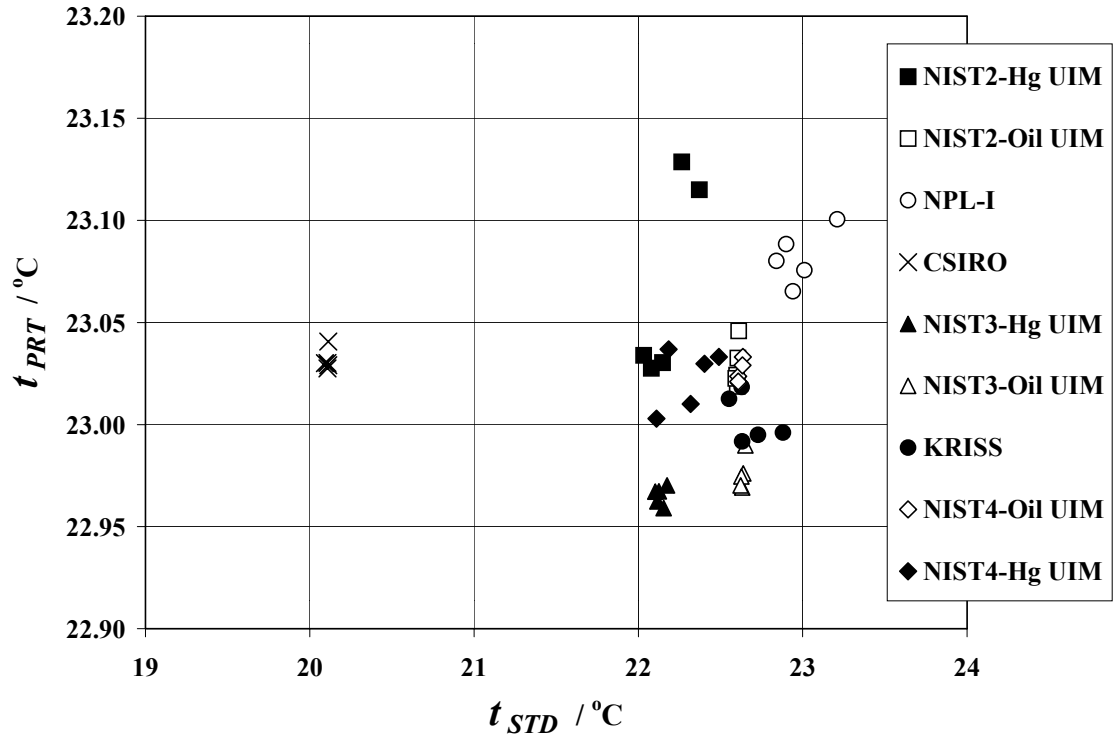
where P_{Sjkl} is the uncorrected pressure measured/generated by the primary standard operating at absolute temperature T_{Sjkl}, T_{Gikl} is the absolute temperature of the transfer standard gauge, *a*, *b*, and *c*, are temperature-independent gas-species dependent constants, and the parameter *Y* is defined as

$$Y = 2P_{Sjkl} d \Big/ \left[133\left(T_{Sjkl} + T_{Gikl}\right)\right] \qquad (2b)$$

and *d* is the internal diameter of the gauge inlet tubing in *mm* if pressure is in *Pa*. The interior temperature of the enclosure as measured by the PRT is assumed to closely approximate the temperature of each gauge. It can be seen in Figure 2 that the temperature of the transfer standards was only slightly affected (~ 0.1 ^{o}C) by relatively large differences in laboratory temperatures at different NMIs, as indicated by the differences in operating temperatures of the primary standards (19.5 ^{o}C to 23.5 ^{o}C).



Transfer Package A



Transfer Package B

Figure 2. Temperature of the transfer standard (t_{PRT}) as a function of primary standard temperature (t_{STD}) at the beginning of each calibration run.

The corrections to the measured/generated pressures (factors f_{tt}) are based on d = 4.6 mm and the following reported values for nitrogen: $a = 1.2 \times 10^6$, $b = 1.0 \times 10^3$, and $c = 14$ [23]. The corrections, which are relatively small, are largest at lowest pressures. For example at 1 Pa, typical values for f_{tt} in Equation (2a) are: 1.0003 (NIST), 1.0012 to 1.0017 (PTB), 1.0041 (NPL-UK), 1.0033 to 1.0043 (IMGC), 0.9997 to 1.0003 (NPL-I), and 1.0038 (CSIRO).

6.3. CALCULATION OF CALIBRATION RATIOS

The transfer standard gauges are nominally linear devices and so the ratio of transfer standard reading to primary standard reading will be essentially independent of pressure for a range of pressures about each target value. Once calculated these calibration ratios are used to correct the gauge readings for deviations of the primary standard from the target pressure and so they form the basis for the comparison of measurement standards from different NMIs.

At each target pressure during calibration run k the mean ratio of 5 sets of repeat readings of transfer standard gauge i and primary standard j is given by

$$a_{ijk} = \frac{1}{5} \sum_{l=1}^{5} \frac{p_{ikl}}{P_{jkl}} \tag{3}$$

where p_{ikl} and P_{jkl} are the "simultaneous" readings of the gauge and primary standard, respectively. The mean of the a_{ijk} for 5 calibration runs defines a *calibration ratio* given by

$$a_{ij} = \frac{1}{5}\sum_{k=1}^{5} a_{ijk} = \frac{1}{25}\sum_{k=1}^{5} \sum_{l=1}^{5} \frac{p_{ikl}}{P_{jkl}} \tag{4}$$

The calibration ratio, if expressed as

$$a_{ij} = \frac{p_i}{P_j}, \tag{5}$$

may be used to calculate a gauge reading p_i from the pressure being measured/generated[10] by primary standard j, P_j, or vice-versa.

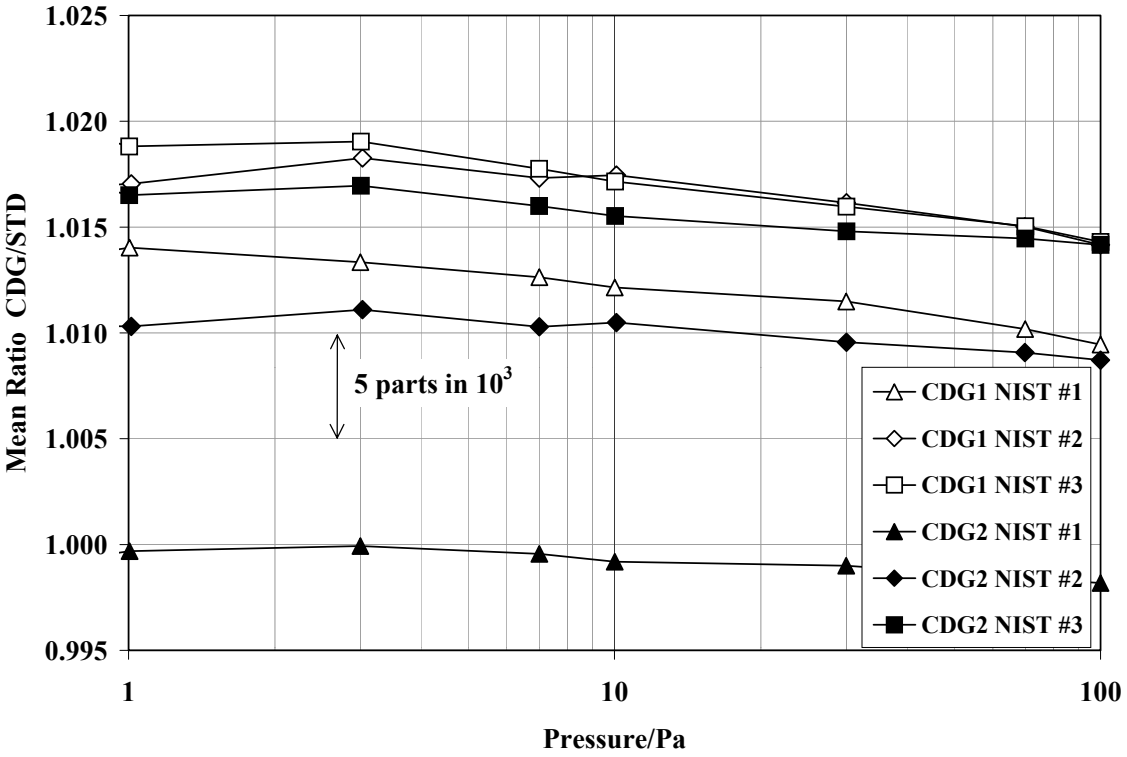
Figures 3 and 4 present the calibration ratios for CDGs and RSGs in the two transfer standard packages as determined by four calibration cycles at NIST, three absolute-mode calibrations of each package. The superior stability of the RSGs is clearly evident, even at 100 Pa where the long-term shifts in their response between calibrations is about a factor of 50 smaller than similar shifts exhibited by the CDGs. It is remarkable that, despite the damage sustained by Package B during shipment to the pilot laboratory for calibration NIST #3, the shifts in calibration of its gauges were not unusual nor significantly different (in magnitude) from calibration shifts observed for gauges in Package A.

6.4. RE-SCALING THE CDG READINGS

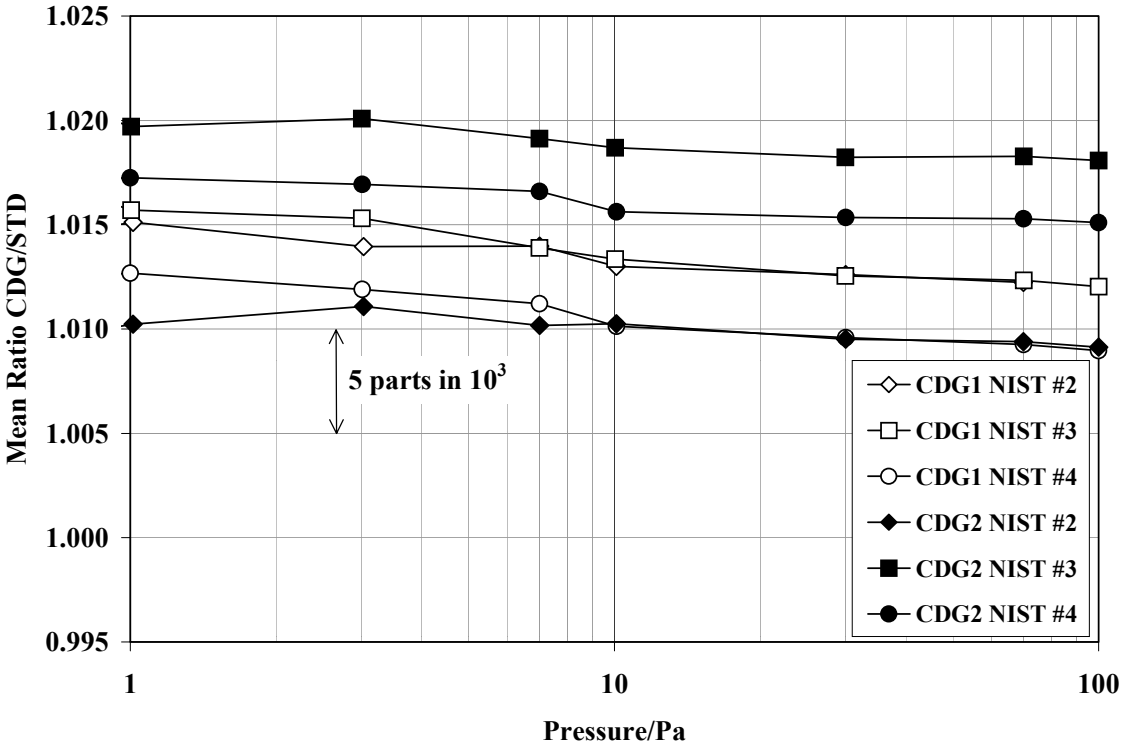
The relatively large calibration shifts of the CDGs can be reduced significantly by re-scaling their readings so they equal those of the RSGs (in the same package) at an overlapping pressure, namely 100 Pa. The readings of the two CDGs at 100 Pa could be re-scaled to a single RSG (either RSG1 or RSG2) or to the mean of two RSGs but then the re-scaled readings of CDG1 and CDG2 would not be independent as required for Youden graphical analyses in Section 7. Although arbitrary, it was decided to pair CDG1 with RSG1, and CDG2 with RSG2, when re-scaling the CDG readings.

---

[10] The measured or generated pressure is the calculated value obtained from measurements of mercury (or oil) temperature and column-height changes in manometers, or from measurements of initial pressure, volume ratios and gas temperatures in expansion systems, respectively.

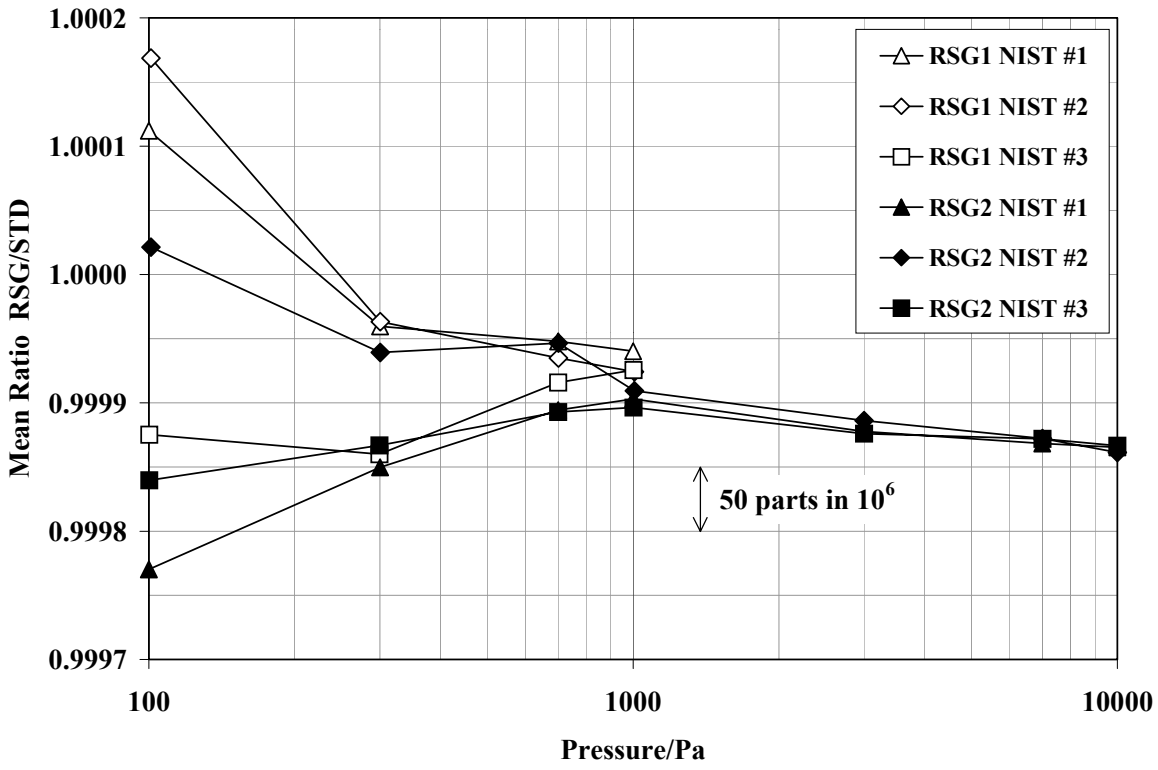


Transfer Package A

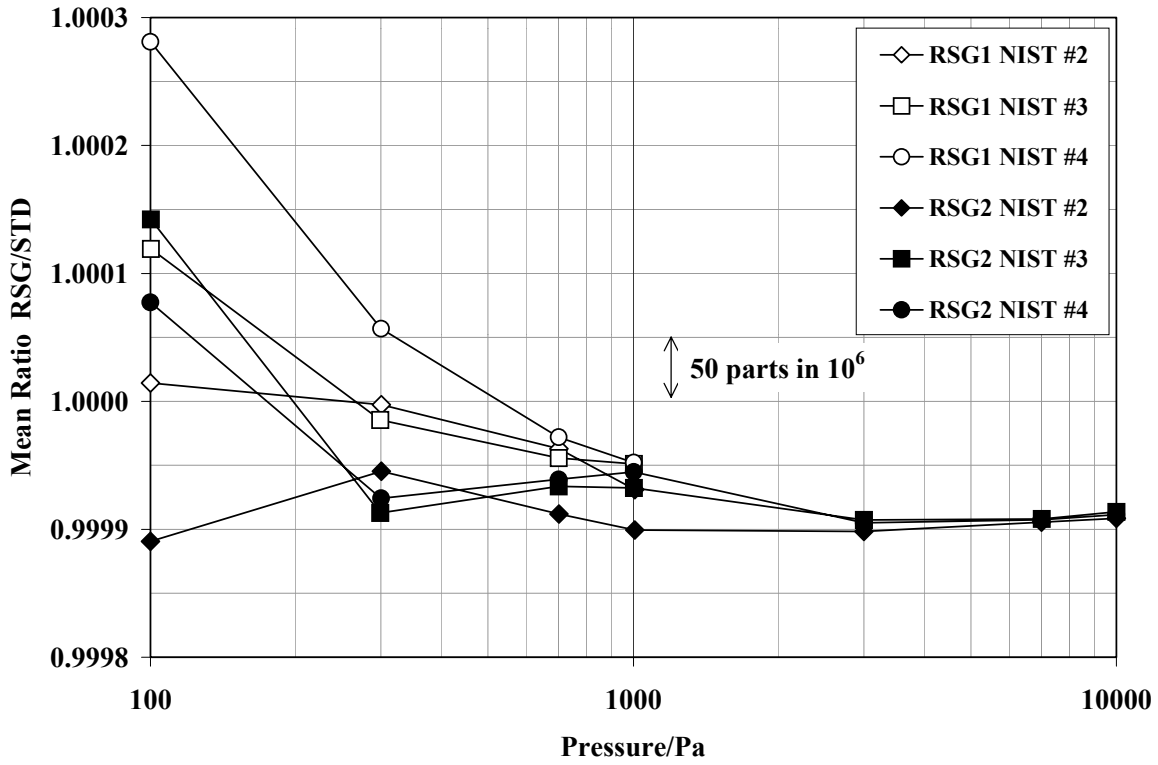


Transfer Package B

Figure 3. Calibration ratios for CDGs in Transfer Package A (upper) and Transfer Package B (lower) as a function of pressure.

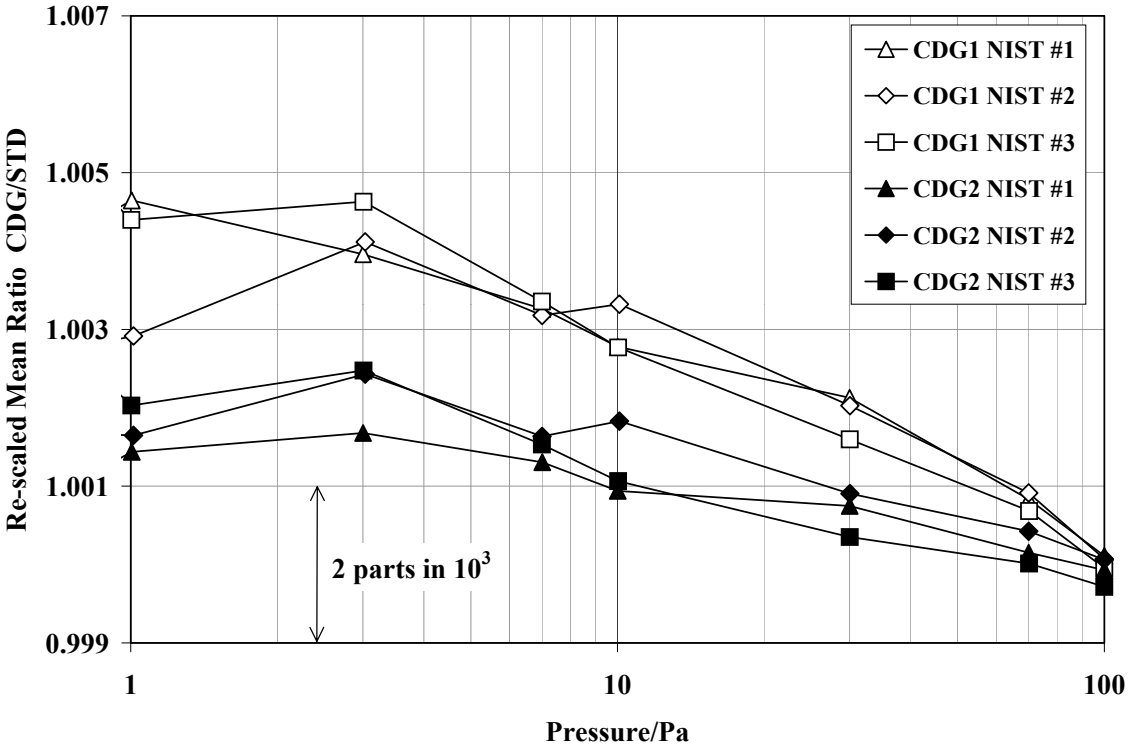


Transfer Package A

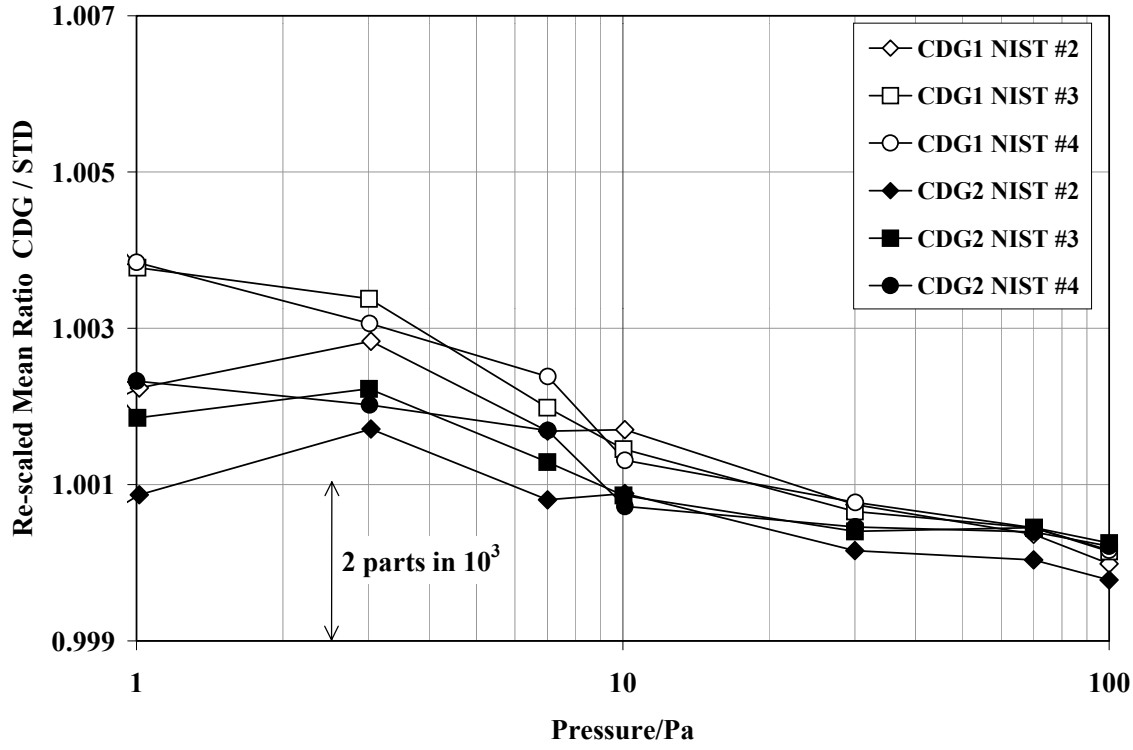


Transfer Package B

Figure 4. Calibration ratios for RSGs in Transfer Package A (upper) and Transfer Package B (lower) as a function of pressure.



Transfer Package A



Transfer Package B

Figure 5. Calibration ratios for the CDGs after re-scaling to the RSGs, as a function of pressure.

At target pressures p_t < 100 Pa, the re-scaled reading of capacitance diaphragm gauge i may be expressed as

$$p_{CDGi}(p_t) = p_{Gi}(p_t)\left[\frac{p_{RSGi}(100)}{p_{Gi}(100)}\right] \tag{6}$$

where $p_{Gi}(p_t)$ is the CDG reading before re-scaling. This equation may be re-expressed in terms of calibration ratios by means of Equation (5) as

$$a_{CDGij}(p_t) = a_{Gij}(p_t)\left[\frac{a_{RSGij}(100)}{a_{Gij}(100)}\right] \tag{7}$$

where a_{Gij} and a_{CDGij} are the respective calibration ratios for capacitance gauge i before and after re-scaling, and a_{RSGij} is the calibration ratio for resonant silicon gauge i. As Figure 5 shows, the observed shifts in the CDG ratios between successive calibrations at NIST are substantially reduced by re-scaling even though sizeable changes in linearity of the CDG response remain at lowest pressures.

In summary, the present key comparison is based on pairs of calibration ratios, a_{CDG1j} and a_{CDG2j} for pressures lower than 100 Pa, and on a_{RSG1j} and a_{RSG2j} for pressures 100 Pa up to and including 1000 Pa.

6.5. Calculation of the Predicted Gauge Readings

Degrees of equivalence [1] of the primary standards for absolute pressure can be expressed quantitatively by comparing pressure readings of the transfer standard gauges. The basic method adopted here is to use the calibration ratios to predict gauge readings that would be observed when different primary standards measure/generate pressures exactly equal to the target value. The difference in the predicted gauge readings is taken as a surrogate for the difference between "true" pressures actually realized by the different primary standards[11]. Results obtained with the two transfer standard packages are normalized by using data taken during simultaneous calibrations at the pilot laboratory.

At target pressures up to and including 1000 Pa there are two gauges (i = 1, 2). Thus for either package there will be two gauge readings for each pressure measured/generated by primary standard j and, according to Equation (5), these may be expressed as:

$$p_{ijA} = a_{ij}P_j \qquad \text{or} \qquad p_{ijB} = b_{ij}P_j \tag{8}$$

where a_{ij} and b_{ij} are the calibration ratios for gauges i in Packages A and B, p_{ijA} and p_{ijB} are their respective pressure readings, and P_j is the measured/generated pressure. Clearly, gauge readings from a given package could be used to compare the primary standards used for their calibration. However to compare primary standards used to calibrate different transfer packages requires that the relationship between gauge readings in the two packages be known.

The relationship between gauge readings i in Packages A and B can be determined by simultaneous calibration against primary standard j and may be expressed as

$$\frac{p_{ijA}}{a_{ij}} = \frac{p_{ijB}}{b_{ij}} = P_j \tag{9}$$

i.e., the ratio of readings of each pair of gauges i in the two packages is equal to the inverse ratio of their calibration ratios determined during the simultaneous calibration. The ratios of calibration ratios, once determined by the simultaneous calibration, could be used to convert all comparison data from Package B to equivalent data from Package A, or vice versa. However, this would in effect reference the key comparison to one package, either Package A or Package B.

---

[11] The difference between "true" pressures being realized by two primary standards, when set to measure/generate a target pressure p_t, is (to a very good approximation) equal to but has the opposite sign of the difference between the measured/generated pressures when both standards realize the same "true" pressure equal to p_t.

Alternatively, Equation (9) can be expressed as

$$\frac{p_{ij}}{n_{iA}a_{ij}} = \frac{p_{ij}}{n_{iB}b_{ij}} = P_j \tag{10}$$

where n_{iA} and n_{iB} are coefficients that re-scale or normalize the readings of gauges i in the two packages to a common normalized gauge reading, p_{ij}, according to

$$p_{ij} = n_{iA}p_{ijA} = n_{iB}p_{ijB} \tag{11}$$

A second property needed to define the normalization coefficients is

$$p_{ij} = (p_{ijA} + p_{ijB})/2 \tag{12}$$

i.e., the normalized gauge reading is also equal to the mean reading of gauges i in Packages A and B obtained during their simultaneous calibration. The significance of Equation (10) when re-written as

$$p_{ij} = n_{iA}a_{ij}P_j = n_{iB}b_{ij}P_j \tag{13}$$

is that it predicts the same normalized gauge reading p_{ij} for a given measured/generated pressure P_j whether calibration ratios for Package A or those for Package B are used. Therefore, once the normalization coefficients have been determined, Equation (13) provides a common basis for comparing results obtained by participants using Package A with results obtained from Package B.

The normalization coefficients, n_{iA} and n_{iB}, were expressed in terms of calibration ratios determined from simultaneous calibration of the two transfer packages at the pilot laboratory (PL) via Equations (11) and (12):

$$n_{iA} = \frac{a_{iPL} + b_{iPL}}{2a_{iPL}} \qquad \text{and} \qquad n_{iB} = \frac{a_{iPL} + b_{iPL}}{2b_{iPL}} \tag{14}$$

Table 2 presents values for the coefficients and their mean obtained from two simultaneous calibrations, NIST #2 and #3.

Thus, the procedure is to calculate the normalized gauge readings, p_{ij}, that would be obtained when the pressure measured/generated by each primary standard j equals the target pressure, p_t, i.e., when $P_j = p_t$. The respective normalized gauge readings for Package A and for Package B are then obtained from Equation (13) as:

$$p_{ij} = n_{iA}a_{ij}p_t \qquad \text{and} \qquad p_{ij} = n_{iB}b_{ij}p_t \tag{15}$$

The results for p_{ij} from individual laboratories are presented in Section 7.1.

At each target pressure up to and including 1000 Pa, there are two values for the normalized gauge reading (e.g., for CDG1 and CDG2, etc.) and so a mean normalized gauge reading p_{jU} was calculated as a simple arithmetic mean:

$$p_{jU} = \frac{p_{1j} + p_{2j}}{2} \tag{16}$$

where the subscript U denotes that gauge readings p_{ij} are uncorrected. For the pilot laboratory, a single value of p_{jU} was calculated as the arithmetic mean of twelve values of p_{ij}^{mn}. The values of p_{ij}^{mn} were determined via Equation (15) using calibration ratios a_{ij}^{mn} or b_{ij}^{mn} obtained from three calibrations (n = 1, 2, 3) of two gauges (i = 1, 2) in two transfer standard packages (m = A, B).

Table 2. Normalization coefficients n_{iA} and n_{iB} for gauges in transfer standard packages A and B, respectively, where values above the dotted line refer to CDGs and those below refer to RSGs.

Package	Target Pressure Pa	From NIST #2		From NIST #3		Mean Coefficient	
		CDG1	CDG2	CDG1	CDG2	CDG1	CDG2
		RSG1	RSG2	RSG1	RSG2	RSG1	RSG2
A	1	0.999662	0.999611	0.999580	0.999802	0.999621	0.999707
	3	0.999363	0.999642	0.999297	0.999795	0.999330	0.999718
	10	0.999194	0.999529	0.999311	0.999866	0.999252	0.999698
	30	0.999361	0.999625	0.999524	1.000017	0.999442	0.999821
	100	0.999957	0.999858	1.000094	1.000269	1.000026	1.000063
	300	1.000017	1.000003	1.000063	1.000023	1.000040	1.000013
	1000	1.000003	0.999995	1.000013	1.000018	1.000008	1.000006
B	1	1.000338	1.000389	1.000421	1.000198	1.000379	1.000293
	3	1.000638	1.000359	1.000704	1.000205	1.000671	1.000282
	10	1.000807	1.000471	1.000690	1.000134	1.000749	1.000303
	30	1.000640	1.000375	1.000477	0.999983	1.000558	1.000179
	100	1.000043	1.000142	0.999906	0.999732	0.999974	0.999937
	300	0.999983	0.999997	0.999937	0.999977	0.999960	0.999987
	1000	0.999997	1.000005	0.999987	0.999982	0.999992	0.999994

The "true" pressures realized by the primary standards when set to measure/generate a given target pressure should, on average, closely approximate the SI value under the assumption that deviations from the SI value are randomly distributed. Therefore, it is reasonable to correct the mean normalized gauge readings so that their ensemble average is also equal to the target pressure. Thus, the corrected mean gauge reading can be expressed as

$$p_j = f_C p_{jU} = f_C \left(\frac{p_{1j} + p_{2j}}{2} \right) \tag{17}$$

where the correction factor f_C is obtained by setting the ensemble average of the p_j values from calibrations at individual laboratories equal to the target pressure (see Section A1 of the Appendix). The resultant values for f_C are very nearly equal to one (see Table A1). The results for p_j from individual laboratories are presented in Section 7.2.

Implicit in the above analysis is the assumption that response functions of the transfer standard gauges do not change during the comparison. Of course this is not true (see Figures 4 and 5) since the long-term shifts in gauge response as well as other sources will contribute uncertainty to the normalized gauge readings p_{ij} and ultimately to the corrected mean gauge readings p_j.

6.6. Estimates of Uncertainties in the Normalized Gauge Readings

The combined uncertainty[12] in the normalized gauge readings calculated using Equation (15) may be estimated from the root-sum-square of three component uncertainties [25, 26],

$$u_c(p_{ij}) = \sqrt{u_{std}^2(p_{ij}) + u_{rdm}^2(p_{ij}) + u_{lts}^2(p_{ij})} \tag{18}$$

where $u_{std}(p_{ij})$ is the uncertainty in p_{ij} due to systematic effects in primary standard j, $u_{rdm}(p_{ij})$ is the uncertainty in p_{ij} due to the combined effect of short-term random errors of transfer standard gauge i and

[12] Uncertainty refers to standard uncertainty unless noted otherwise.

primary standard j during calibration, and $u_{lts}(p_{ij})$ is the uncertainty arising from long-term shifts in the response function of gauge i during the course of the comparison.

Table 3 and Figure 6 present the estimated relative uncertainties in pressure arising from systematic effects in the primary standards, as stated by the participants for target pressures used in the comparison. Such estimates usually involve both Type A and Type B evaluations. It is evident from the plot that relative uncertainties for static expansion systems are only weakly dependent on pressure whereas those for liquid-column manometers decrease rapidly with increasing pressure. At 1000 Pa the uncertainties for the manometers are one to two orders of magnitude lower than those stated for expansion systems.

The relative uncertainty in p_{ij} due to short-term random effects during calibration can be estimated from the corresponding uncertainties in the calibration ratios via Equation (15):

$$\frac{u_{rdm}(p_{ij})}{p_{ij}} = \frac{u_{rdm}(a_{ij})}{a_{ij}} \tag{19}$$

Similarly the relative uncertainty in p_{ij} due to long-term shifts in gauge response between calibrations is given by

$$\frac{u_{lts}(p_{ij})}{p_{ij}} = \frac{u_{lts}(a_{ij}^{'})}{a_{ij}^{'}} \tag{20}$$

where $a_{ij}^{'} = n_{iA} a_{ij}$ is the normalized calibration ratio for gauge i in Package A. It is understood that the above equations involving a_{ij} or n_{iA} apply equally well to b_{ij} and n_{iB}.

The short-term random uncertainty in a calibration ratio, a_{ij}, as given by Equation (4), may be estimated by a Type A evaluation in one of two ways, either as

$$u_{rdm}(a_{ij}) = \sigma_{ijkl} \Big/ \sqrt{25} \tag{21}$$

where σ_{ijkl} is the standard deviation of 25 values of p_{ikl}/P_{jkl} about their mean a_{ij} OR

$$u_{rdm}(a_{ij}) = \sigma_{ijk} \Big/ \sqrt{5} \tag{22}$$

where σ_{ijk} is the standard deviation of five values of the means, a_{ijk}, about their mean a_{ij}. Equation (21) is appropriate only if the "true" means for the five runs are equal, that is, they refer to the same parent population of observations. If not, it suggests that the run-to-run (day-to-day) variability is dominant in which case Equation (22) should be used to calculate short-term random uncertainties.

An analysis of variance (ANOVA) was performed on the five sets of five pressure ratios for each gauge at each target pressure to test the hypothesis that the "true" means from five calibration runs are equal. For nearly all comparison data, the hypothesis was rejected with a less than 5 % chance that the means are actually equal. For the few remaining data, the hypothesis was rejected with a somewhat larger chance for error (~50 % or less). Therefore, Equation (22) was used to estimate the short-term random uncertainties, $u_{rdm}(a_{ij})$ or $u_{rdm}(b_{ij})$, which are given in columns seven and eight of Table 5 in Section 7.1. The short-term random uncertainties in the re-scaled calibration ratios obtained via Equation (7) were estimated as the root-sum-square of component uncertainties arising from random effects in $a_{Gij}(p_t)$, $a_{Gij}(100)$, and $a_{RSGij}(100)$, each evaluated using Equation (22).

The ANOVA tests confirm an expected result, and that is, random effects due to operational differences between five calibration runs performed on five different days are generally not the same as, and are larger than, random effects associated with five repeat readings taken during a period of five minutes. Possible sources of the short-term day-to-day variability include differences in zero drifts of the gauges, differences in achieving stable target pressures, etc. This variability was assumed to be random and uncorrelated for each pair of gauges.

Table 3. Relative standard uncertainties, as stated by the participants, due to systematic effects in their primary standards. Pilot laboratory values above the dashed line refer to an oil UIM and those below to a mercury UIM. Not all digits are significant but are retained for calculation of final results.

Target Pressure	$100 \times u_{std}(p_t)/p_t$							
	Static Expansion Systems				Liquid-Column Manometers			
Pa	IMGC	NPL-I	NPL-UK	PTB	CSIRO	IMGC	KRISS	NIST
1	0.3	0.0715	0.3015	0.134	1.702			0.1011
3	0.3	0.055	0.3005	0.122	0.572			0.0500
10	0.3	0.055	0.3002	0.122	0.186		0.1301	0.0151
30	0.3	0.055	0.1893	0.106	0.0940		0.0436	0.0053
100	0.3	0.052	0.1828	0.106	0.0769	0.0372	0.0139	0.0023
300		0.052	0.1809	0.106	0.0500	0.0126	0.0066	0.0010
1000		0.073	0.1803	0.111	0.0150	0.0039	0.0052	0.0004

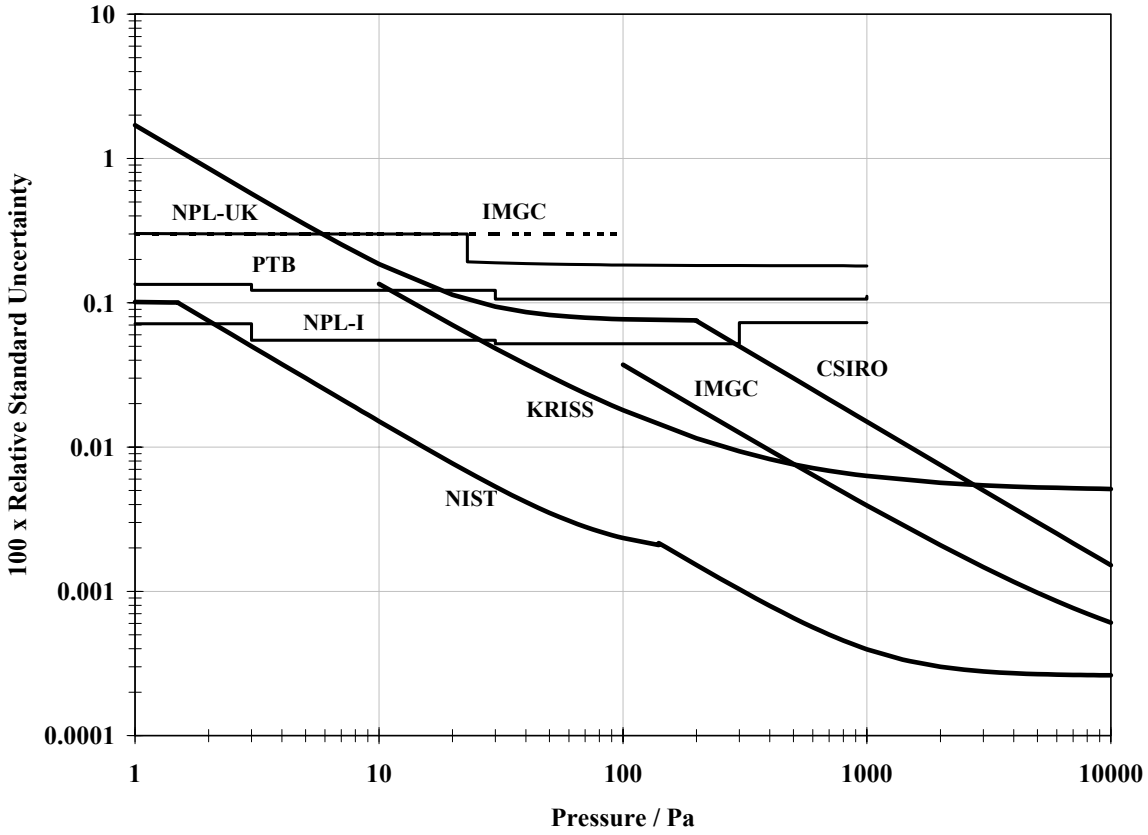


Figure 6. Relative standard uncertainty due to systematic effects in primary standards at the participating laboratories as a function of pressure. The heavy lines identify uncertainties associated with liquid-column manometers and the medium solid/dashed lines identify uncertainties for the static expansion systems.

Long-term shifts in gauge response are often one of the largest component uncertainties, particularly for CDGs, yet most difficult to evaluate. The reasons for this are often twofold: (a) the number of repeat calibrations against the same standard is limited for practical reasons and (b) the effect of transportation between laboratories (rough handling, etc.) is unknown. Earlier studies at the pilot laboratory [21, 22] have shown that changes in response functions of CDGs and RSGs between calibrations generally do not occur as a monotonic drift with time (over intervals of months to years) but rather as shifts that are essentially random in both sign and magnitude. Furthermore, the earlier studies showed that, at least for low range CDGs, the magnitude of the shifts was on average about a factor of two larger for gauges transported between laboratories than for gauges maintained at the pilot laboratory.

In the present comparison, the observed shifts in gauge response between calibrations at the pilot laboratory (see Figures 3 and 4) are consistent with the earlier studies, i.e., there is little evidence that the calibration shifts are systematic (even for the RSGs) but rather they appear more random in character. Therefore, the observed variability in gauge response was assumed to be purely random but, because the statistical sample of pilot laboratory calibrations was limited (three), a Type B evaluation was used to estimate the uncertainty $u_{lts}(a_{ij})$ for each gauge.

At a given target pressure, the variation due to long-term shifts was modeled by a normal distribution such that the best estimated value is $((a_{i\,PL})_{\max} + (a_{i\,PL})_{\min})/2$ and there is a 2 out of 3 chance the calibration ratio lies in the interval between maximum and minimum values of a_{iPL} obtained from three calibrations at the pilot laboratory. Then the standard uncertainty due to this source of error equals one-half the difference between the maximum and minimum values:

$$u_{lts}(a_{ij}) = \left((a_{iPL})_{\max} - (a_{iPL})_{\min}\right)/2 \tag{23}$$

This estimate is unaffected by any systematic bias in the pilot laboratory primary standard, which would be present in all three calibrations. However the estimate does assume that the observed shifts in the calibration ratios are primarily due to the gauges and not the primary standard at the pilot laboratory.

In order to check the latter assumption, results shown in Figures 3 and 4 for the pilot laboratory calibrations are combined in Figure 7 where calibration ratios for gauge 2 are plotted against those for gauge 1. The results correspond to calibrations of CDGs against the oil UIM (upper graph) at target pressures of 1 Pa, 3 Pa, 7 Pa, 10 Pa, 30 Pa, 70 Pa, and 100 Pa, and calibrations of RSGs against the mercury UIM (lower graph) at target pressures of 100 Pa, 300 Pa, 700 Pa and 1000 Pa. Individual points for a given calibration represent results for a pair of gauges at different target pressures where the black symbols refer to Package A and the gray symbols refer to Package B. Dashed lines connect the results for three calibrations of CDGs at 10 Pa and RSGs at 300 Pa in Package A (white + symbols) and Package B (black + symbols). A correlation is seen between CDG2 and CDG1 for a given calibration but this is due to rather similar non-linear behavior of the gauges as a function of pressure. If the observed shifts were due to a common source, such as the pilot laboratory primary standard, they would have the same sign and magnitude for each pair of gauges at a given target pressure (all dashed lines would be aligned diagonally). As may be seen, there is essentially no evidence of a correlation between calibrations at a given target pressure for the CDGs and little evidence of correlation between the RSGs, which are significantly more stable. As shown later in this section (see Figure 8) the long-term instabilities of the RSGs and the re-scaled CDGs are of the same order as, or larger than, the uncertainty due to systematic effects in the primary standards at the pilot laboratory. Several UIMs developed at the pilot laboratory including the UIMs used in this comparison have been checked by direct comparison and their stability found to be consistent with their stated uncertainties. Therefore in the absence of definitive evidence to the contrary, the statement that the observed shifts in the calibration ratios are primarily due to the gauges is taken to be a valid assumption for this key comparison.

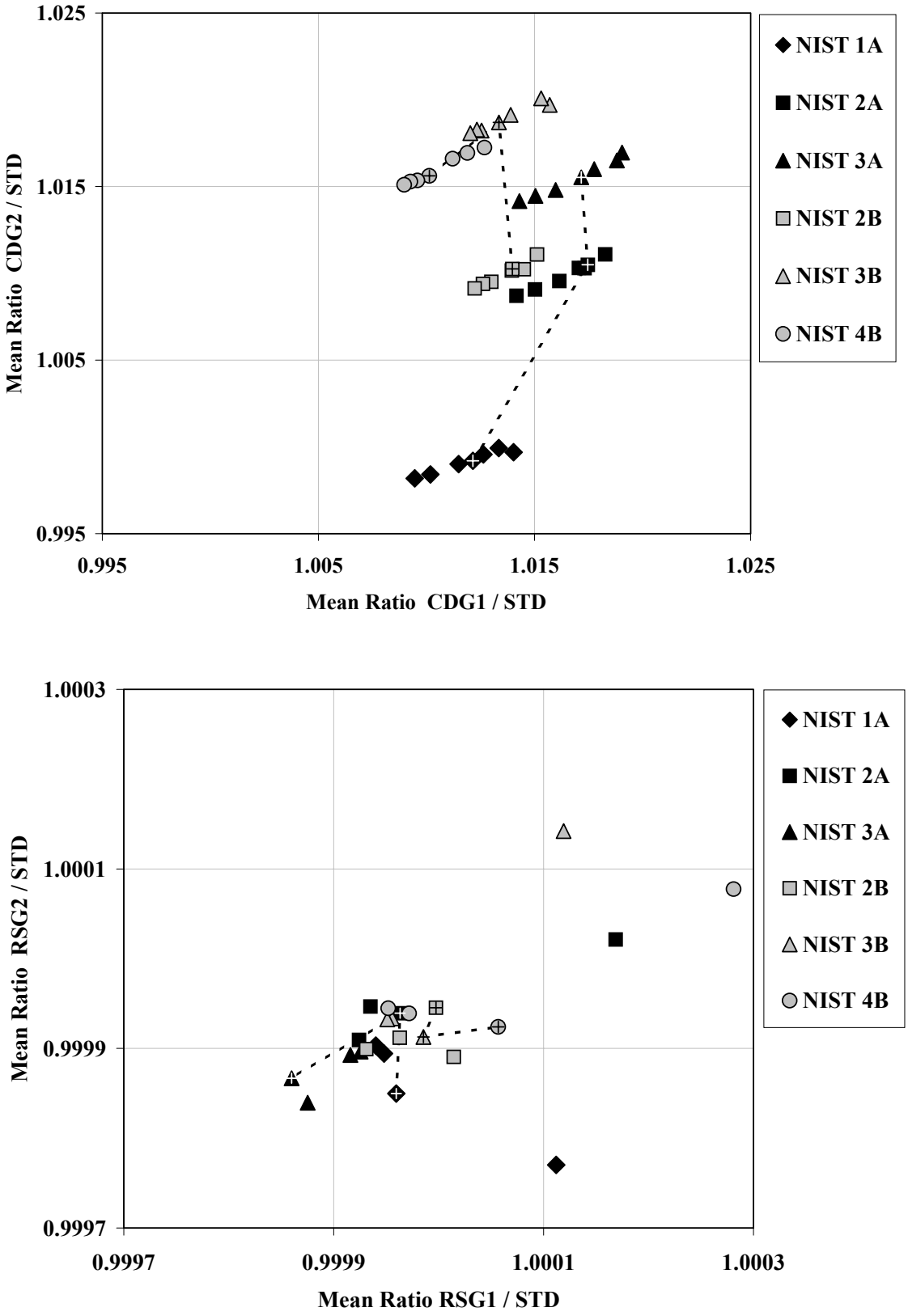


Figure 7. Calibration ratios from pilot laboratory data for gauge 2 versus those for gauge 1. Individual points refer to data at different target pressures. Dashed lines connect results for CDGs at 10 Pa and for RSGs at 300 Pa.

The possibility exists that long-term shifts associated with Equation (23) and short-term random effects associated with Equation (22) refer to the same source of variability. If true, the uncertainty contributions from short-term and long-term variability in a_{ij} would not be independent and should be included only once in estimating the combined uncertainty in the normalized gauge readings. As a check, an ANOVA was performed on the 15 mean ratios obtained from three calibrations at the pilot laboratory for each gauge and at each target pressure to test the hypothesis that the "true" means of the three groups of five ratios are equal. For nearly all pilot laboratory data on CDGs, the hypotheses were rejected with a less than 5 % chance that the means are actually equal. For RSGs the hypotheses were rejected with a less than 25 % chance on average that the means are equal. Although there was a non-negligible chance of equal means for the RSGs (due to their superior calibration stability) it is important that an underestimation of uncertainty in the results be avoided. Therefore contributions from both long-term shifts and short-term random errors in the a_{ij}, for the RSGs as well as the CDGs, were included in estimates of combined uncertainty in the normalized pressure readings of the gauges. The variability due to long-term shifts in gauge response was assumed to be random and uncorrelated for each pair of gauges and for the calibrations at different laboratories.

The normalized calibration ratio, $a_{ij}^{'} = n_{iA}\, a_{ij} = n_{iA}(a_{iPL2}, b_{iPL2}, a_{iPL3}, b_{iPL3})\, a_{ij}$, is not only a function of the calibration ratio determined by primary standard j but, through the mean normalization coefficient[13], is also a function of the calibration ratios a_{iPL} and b_{iPL} obtained from two simultaneous calibrations at the pilot laboratory (NIST #2 and #3). Taking into account the propagation of uncertainties, the uncertainty in the normalized calibration coefficients for Package A due to long-term shifts may be estimated from

$$u_{lts}^2(a_{ij}^{'}) \simeq n_{iA}^2\, u_{lts}^2(a_{ij}) + (1/8)\, u_{lts}^2(a_{ij}) + (1/8)\, u_{lts}^2(b_{ij}) - r_{jPL}\,(n_{iA}/2)\, u_{lts}^2(a_{ij}) \qquad (24)$$

where $u_{lts}(a_{iPL2}) = u_{lts}(a_{iPL3}) = u_{lts}(a_{ij})$ and $u_{lts}(b_{iPL2}) = u_{lts}(b_{iPL3}) = u_{lts}(b_{ij})$. The partial derivatives (e.g., $\partial a_{ij}^{'}/\partial a_{iPL2}$, etc.) were evaluated using the following approximations: $a_{iPL2} \simeq a_{iPL3} \simeq a_{ij}$, $b_{iPL2} \simeq b_{iPL3} \simeq b_{ij}$, and $(b_{iPL}/a_{iPL}) \approx 1$. The correlation coefficient r_{jPL} is equal to unity when j is either *PL*2 or *PL*3 (NIST #2 or NIST #3) but is zero for all other calibrations of Package A. Similarly for Package B:

$$u_{lts}^2(b_{ij}^{'}) \simeq n_{iB}^2\, u_{lts}^2(b_{ij}) + (1/8)\, u_{lts}^2(b_{ij}) + (1/8)\, u_{lts}^2(a_{ij}) - r_{jPL}\,(n_{iB}/2)\, u_{lts}^2(b_{ij}) \qquad (25)$$

The relative uncertainties in the calibration ratios due to long-term shifts in gauge response and their normalized counterparts were estimated using Equations (23) to (25) and are given in Table 4. The estimates for CDGs are based on variability of their calibration ratios after re-scaling to the RSGs.

It is noteworthy that the relative uncertainty in the normalized gauge readings [equal to the relative uncertainty in the corresponding normalized ratio via Equation (20)] is of the same order of magnitude as, or in some cases smaller than, the relative uncertainty due to systematic effects in the primary standards. This is illustrated in Figure 8 where the relative uncertainties of the transfer standard gauge readings are superimposed upon the relative uncertainties of the primary standards (shown in Figure 6). This plot shows that the long-term stability of the transfer standard over the course of this comparison should be sufficient to resolve any relative biases between different primary standards.

Finally, the combined uncertainty in the normalized gauge readings, p_{ij}, at each target pressure was estimated by using data from Tables 3, 4, and 5 and Equations (18) to (20), and is given in Table 5.

---

[13] The mean is derived from Equation (14) as $n_{iA} = (1/2)\left[\left(\frac{a_{iPL2} + b_{iPL2}}{2a_{iPL2}}\right) + \left(\frac{a_{iPL3} + b_{iPL3}}{2a_{iPL3}}\right)\right] = (1/4)\left[2 + \left(\frac{b_{iPL2}}{a_{PL2}}\right) + \left(\frac{b_{iPL3}}{a_{PL3}}\right)\right]$

Table 4. Relative standard uncertainty in the calibration ratio a_{ij} or b_{ij} and in the normalized calibration ratio a'_{ij} or b'_{ij} due to long-term shifts in gauge response. The values in parentheses refer to the uncertainty in the normalized ratios obtained from either of the two simultaneous calibrations at the pilot laboratory, NIST #2 or NIST #3. Values above the dotted line refer to CDGs and those below refer to RSGs. Not all digits are significant but are retained for calculation of final results.

Target	$100\times u_{lts}(a_{ij})/a_{ij}$		$100\times u_{lts}(b_{ij})/b_{ij}$		$100\times u_{lts}(a'_{ij})/a'_{ij}$				$100\times u_{lts}(b'_{ij})/b'_{ij}$			
Press.	CDG1	CDG2	CDG1	CDG2	CDG1		CDG2		CDG1		CDG2	
Pa	RSG1	RSG2	RSG1	RSG2	RSG1		RSG2		RSG1		RSG2	
1	0.0861	0.0407	0.0801	0.0728	0.0962	(0.0737)	0.0455	(0.0412)	0.0903	(0.0703)	0.0785	(0.0593)
3	0.0415	0.0479	0.0271	0.0257	0.0464	(0.0342)	0.0536	(0.0390)	0.0323	(0.0260)	0.0321	(0.0265)
10	0.0272	0.0444	0.0196	0.0083	0.0304	(0.0226)	0.0496	(0.0352)	0.0229	(0.0182)	0.0180	(0.0170)
30	0.0257	0.0269	0.0058	0.0151	0.0287	(0.0204)	0.0301	(0.0220)	0.0110	(0.0102)	0.0187	(0.0153)
100	0.0078	0.0173	0.0087	0.0238	0.0088	(0.0069)	0.0201	(0.0160)	0.0096	(0.0074)	0.0260	(0.0198)
300	0.0052	0.0045	0.0036	0.0016	0.0056	(0.0043)	0.0048	(0.0036)	0.0042	(0.0034)	0.0023	(0.0020)
1000	0.00080	0.00065	0.00105	0.00227	0.00093	(0.00073)	0.00106	(0.00095)	0.00115	(0.00088)	0.00242	(0.00181)

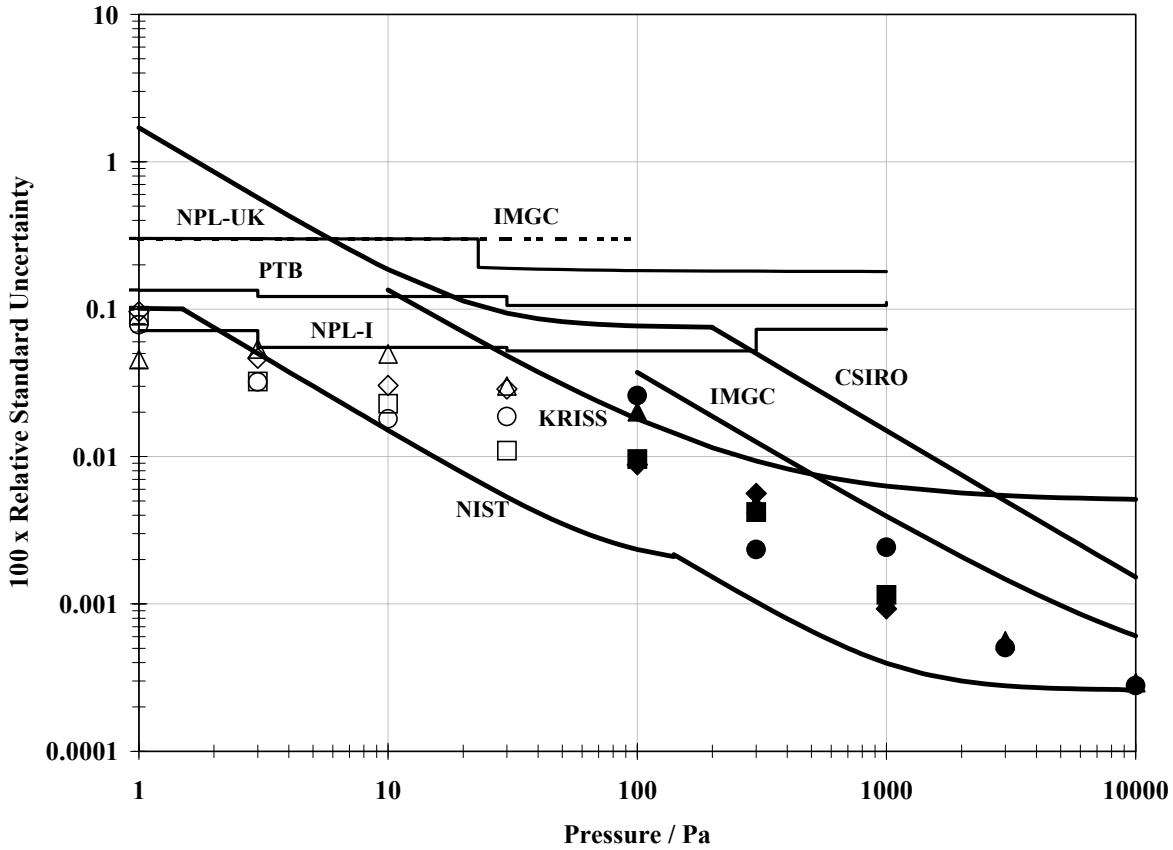


Figure 8. Comparison of the relative uncertainty due to long-term shifts in the transfer standard gauges with the relative uncertainties due to systematic effects in the primary standards. The solid symbols refer to RSGs, the open symbols to (re-scaled) CDGs. Diamond and triangle symbols refer to gauges 1 and 2 in Package A. Square and circular symbols refer to gauges 1 and 2 in Package B.

6.7. ESTIMATES OF UNCERTAINTIES IN THE CORRECTED MEAN GAUGE READINGS

The component uncertainties in $u_c(p_{ij})$ will also propagate to the combined uncertainty in the corrected mean gauge reading p_j calculated via Equation (17). For the non-pilot laboratories, the combined uncertainty was estimated from [25, 26]

$$u_c^2(p_j) \simeq u_c^2(p_{jU}) = u_{std}^2(p_j) + \sum_{i=1}^{2} c^2 u_{rdm}^2(p_{ij}) + \sum_{i=1}^{2} c^2 u_{lts}^2(p_{ij}) \tag{26}$$

where $u_{std}(p_{1j}) = u_{std}(p_{2j}) = u_{std}(p_{jU}) = u_{std}(p_j)$, $c = ½$ is the (common) value for the partial derivatives, $\partial p_{jU} / \partial p_{ij}$, and the approximation $f_C \simeq 1$ was used.

For the pilot laboratory, p_{jU} is the mean of twelve values of p_{ij}^{mn} at target pressures up to and including 1000 Pa, where m is the package label and n is the calibration number [see discussion following Equation (16)]. In this case the combined uncertainty in p_j was estimated from:

$$u_c^2(p_j) \simeq u_c^2(p_{jU}) = u_{std}^2(p_j) + \sum_{m=A}^{B} \sum_{n=1}^{3} \sum_{i=1}^{2} c^2 u_{rdm}^2(p_{ij}^{mn}) + \sum_{m=A}^{B} \sum_{n=1}^{3} \sum_{i=1}^{2} c^2 u_{lts}^2(p_{ij}^{mn}) \tag{27}$$

where $c = 1/12$. Note that multiple calibrations at the pilot laboratory tend to reduce the influence of uncorrelated uncertainties arising from short-term and long-term variability of the gauges on the combined uncertainty in p_j for the pilot laboratory.

6.8. EVALUATION OF DEGREES OF EQUIVALENCE

The Mutual Recognition Arrangement (MRA) [1] proposes that the degree of equivalence of a national measurement standard may be stated in two ways, degree of equivalence relative to a key comparison reference value (KCRV) and degree of equivalence between pairs of national standards. Several procedures can be used to define a KCRV each having both advantages and disadvantages, as described in the Appendix. The definition of a KCRV at each target pressure is given by Equation (A1) of the Appendix, which in effect sets the reference value numerically equal to the target pressure.

The degree of equivalence of primary standard j relative to a KCRV is expressed at each target pressure by two quantities, the deviation of p_j from the reference value p_R

$$D_j = p_j - p_R \tag{28}$$

and the expanded uncertainty of this deviation, which is estimated from

$$U_j^2 = k_{95}^2 u_c^2(D_j) = k_{95}^2 \left[\left(1 - 1/N\right) u_c^2(p_j) + u_c^2(p_R) \right] \tag{29}$$

where $u_c(D_j)$ is the combined standard uncertainty of this deviation, k_{95} is the coverage factor that approximates a 95 % level of confidence for the interval defined by U_j, $u_c(p_j)$ and $u_c(p_R)$ are the combined uncertainties in the corrected mean gauge readings and the reference value given by Equations (27) or (28) and (A5), respectively, and N is the number of primary standards of a given type and is equal to either N_{MAN} or N_{SES} depending on whether primary standard j is a manometer or an expansion system. The term involving $-1/N$ is a correction for the correlation between p_R and p_j.

Following the wording given in the MRA, the degree of equivalence between pairs of primary standards j and j' may be expressed at each target pressure by two quantities, the difference of their deviations from the reference value[14]

$$D_{jj'} = D_j - D_{j'} = (p_j - p_R) - (p_{j'} - p_R) = p_j - p_{j'} \tag{30}$$

[14] The degree of equivalence between pairs of standards is written as stated in the MRA but in reality the difference $D_{jj'}$ does not require the calculation of a KCRV.

and the expanded uncertainty of this difference, which is estimated from

$$U_{jj'}^2 = k_{95}^2 u_c^2(D_{jj'}) = k_{95}^2[u_c^2(p_j) + u_c^2(p_{j'})] \quad (31)$$

where $u_c(D_{jj'})$ is the combined standard uncertainty of this difference, k_{95} is the coverage factor that approximates a 95 % level of confidence for the interval defined by $U_{jj'}$, $u_c(p_j)$ and $u_c(p_{j'})$ are the combined uncertainties in the corrected mean gauge readings obtained with primary standards j and j', respectively, which are estimated from Equation (26) or from Equations (26) and (27).

Values for coverage factors k_{95} that produce the expanded uncertainties U_j and $U_{jj'}$ were obtained using a conventional procedure described in Section A2 of the Appendix and are given in Table A2.

7. RESULTS FOR KEY COMPARISON CCM.P-K4

7.1. COMPARISON OF NORMALIZED GAUGE READINGS

Table 5 presents a summary of the normalized gauge readings, p_{ij}, obtained from calibrations at the pilot and other participant laboratories as a function of nominal target pressures. Results obtained with Package A or with Package B are delineated in the table by a heavy separator line and are presented in chronological order of the calibrations. The calibration ratios for the CDGs, before and after re-scaling to the RSGs, were calculated using Equations (4) and (7), respectively, and are given in column three/four and five/six. The ratios for the RSGs calculated via Equation (4) are also given in column five/six. Uncertainties in the ratios due to short-term random effects, which were obtained by means of Equation (22), are given in column seven/eight. Values of p_{ij}, which were calculated via Equation (15), are given in column nine/ten. The combined standard uncertainties, $u_c(p_{ij})$, which were calculated according to Equation (18), are given in column eleven/twelve.

The results for the normalized gauge readings, p_{ij}, and their standard ($k = 1$) uncertainties, $u_c(p_{ij})$, are presented in Figures 9 through 15 in the form of Youden plots [27] in which the difference $p_{2j} - p_t$ is plotted as a function of $p_{1j} - p_t$. The *y*- and *x*-axes are labeled as *CDG2 - STD* and *CDG1 - STD* or as *RSG2 - STD* and *RSG1 - STD* for greater clarity. Residual errors associated with normalizing gauge readings from different packages to a common reading manifest themselves as differences between the normalized results obtained from two simultaneous calibrations of the two packages at the pilot laboratory. These differences, although small, can be seen in Figures 9 through 12 as differences between NIST 2A and NIST 2B results (black and gray square symbols) and between NIST 3A and NIST 3B results (black and gray triangle symbols). The results from Figures 9, 10, and 11 (upper) are re-scaled and re-plotted in Figures 13, 14 and 15 to include outliers from one of the participant laboratories.

Table 5. Summary of key comparison results for calibration ratios, a_{ij} and b_{ij}, their uncertainty due to short-term random effects, $u_{rdm}(a_{ij})$ and $u_{rdm}(b_{ij})$, calculated values for normalized readings of gauge i, p_{ij}, when the pressure measured/generated by primary standard j equals the target pressure, and their combined standard uncertainty, $u_c(p_{ij})$. Values above the dotted line refer to CDGs and those below refer to RSGs. The dashed lines differentiate results obtained by the pilot laboratory using either an oil UIM or a mercury UIM. Not all digits are significant but are retained for calculation of final results.

NMI	Target Press.	Calibration Ratios Before Re-scaling		a_{ij} or b_{ij}		u_{rdm} (a_{ij} or b_{ij})		p_{ij} / Pa		$u_c(p_{ij})$ / Pa	
				CDG1	CDG2	CDG1	CDG2	CDG1	CDG2	CDG1	CDG2
	Pa	CDG1	CDG2	RSG1	RSG2	RSG1	RSG2	RSG1	RSG2	RSG1	RSG2
	1	1.01403	0.99969	1.00465	1.00144	0.00039	0.00058	1.0043	1.0011	0.0014	0.0013
(Oil UIM)	3	1.01334	0.99993	1.00396	1.00168	0.00030	0.00019	3.0099	3.0042	0.0022	0.0023
	10	1.01215	0.99920	1.00278	1.00094	0.00010	0.00006	10.0203	10.0064	0.0035	0.0052
NIST #1A	30	1.01149	0.99900	1.00213	1.00075	0.00005	0.00004	30.0470	30.0171	0.0089	0.0093
	100	1.00946	0.99819	1.000111	0.999930	0.000045	0.000038	100.0137	99.9993	0.0101	0.0206
(Hg UIM)	300			0.999960	0.999850	0.000010	0.000024	299.9999	299.9588	0.0174	0.0163
	1000			0.999940	0.999903	0.000005	0.000006	999.9483	999.9095	0.0111	0.0129

NMI	Target Press.	Calibration Ratios Before Re-scaling		a_{ij} or b_{ij}		u_{rdm} (a_{ij} or b_{ij})		p_{ij} / Pa		$u_c(p_{ij})$ / Pa	
				CDG1	CDG2	CDG1	CDG2	CDG1	CDG2	CDG1	CDG2
	Pa	CDG1	CDG2	RSG1	RSG2	RSG1	RSG2	RSG1	RSG2	RSG1	RSG2
	1	1.01679	1.00495	1.00426	1.00167	0.00029	0.00021	1.0039	1.0014	0.0017	0.0014
	3	1.01627	1.00485	1.00374	1.00157	0.00021	0.00025	3.0092	3.0039	0.0040	0.0041
	10	1.01507	1.00409	1.00256	1.00082	0.00017	0.00019	10.0181	10.0051	0.0127	0.0133
PTB	30	1.01484	1.00429	1.00234	1.00101	0.00011	0.00015	30.0533	30.0251	0.0331	0.0334
	100	1.01265	1.00332	1.000167	1.000040	0.000040	0.000105	100.0193	100.0103	0.1064	0.1084
	300			0.999685	0.999709	0.000193	0.000188	299.918	299.917	0.324	0.323
	1000			0.999626	0.999620	0.000085	0.000088	999.634	999.627	1.113	1.114
	1	1.01705	1.01031	1.00292	1.00165	0.00018	0.00010	1.0025	1.0014	0.0013	0.0011
(Oil UIM)	3	1.01826	1.01110	1.00412	1.00243	0.00007	0.00008	3.0103	3.0064	0.0018	0.0019
	10	1.01746	1.01049	1.00332	1.00183	0.00004	0.00007	10.0257	10.0153	0.0028	0.0039
NIST #2A	30	1.01614	1.00956	1.00203	1.00091	0.00003	0.00007	30.0440	30.0218	0.0064	0.0071
	100	1.01416	1.00871	1.000072	1.000063	0.000032	0.000066	100.0098	100.0126	0.0079	0.0175
(Hg UIM)	300			0.999963	0.999939	0.000029	0.000033	300.0009	299.9857	0.0158	0.0150
	1000			0.999924	0.999909	0.000010	0.000007	999.9323	999.9159	0.0132	0.0126
	1	1.01918	1.01171	1.00513	1.00295	0.00009	0.00013	1.0047	1.0027	0.0032	0.0031
	3	1.01839	1.01117	1.00434	1.00241	0.00008	0.00013	3.0110	3.0064	0.0091	0.0092
	10	1.01728	1.01046	1.00325	1.00171	0.00013	0.00016	10.0250	10.0141	0.0302	0.0305
NPL-UK	30	1.01505	1.00863	1.00105	0.99990	0.00011	0.00015	30.0149	29.9915	0.0575	0.0577
	100	1.01326	1.00802	0.999288	0.999291	0.000050	0.000117	99.9314	99.9354	0.183	0.184
	300			0.999289	0.999258	0.000078	0.000076	299.799	299.781	0.544	0.543
	1000			1.000608	1.000565	0.000035	0.000034	1000.616	1000.571	1.803	1.803
	1	1.01825	1.01450	1.00583	1.00444	0.00062	0.00058	1.0054	1.0041	0.0032	0.0031
	3	1.01754	1.01396	1.00513	1.00390	0.00049	0.00042	3.0134	3.0109	0.0092	0.0092
	10	1.01653	1.01337	1.00414	1.00332	0.00048	0.00043	10.0338	10.0301	0.0305	0.0307
IMGC-SES	30	1.01615	1.01332	1.00376	1.00327	0.00038	0.00032	30.0961	30.0927	0.0911	0.0910
	100	1.01398	1.01219	1.001610	1.002155	0.000219	0.000089	100.164	100.222		
	300			1.000475	1.000558	0.000162	0.000032	300.155	300.171		
	1000			1.001477	1.001500	0.000118	0.000016	1001.485	1001.507		
	1	1.04821	1.04372	1.03499	1.03282	0.012013	0.011970	1.0346	1.0325		
	3	1.04339	1.03926	1.03023	1.02840	0.007293	0.007335	3.0886	3.0843		
	10	1.02902	1.02533	1.01604	1.01462	0.002648	0.002646	10.1528	10.1432		
IMGC-Hg5	30	1.02062	1.01736	1.00775	1.00673	0.001010	0.001012	30.2158	30.1966		
	100	1.01465	1.01243	1.001855		0.000334		100.1881		0.0507	
	300			1.000383		0.000146		300.1268		0.0603	
	1000			1.000065		0.000046		1000.0729		0.0609	
	1	1.01904	1.01673	1.00462	1.00225	0.00035	0.00012	1.0042	1.0020	0.0013	0.0011
(Oil UIM)	3	1.01921	1.01712	1.00479	1.00264	0.00015	0.00011	3.0124	3.0071	0.0019	0.0019
	10	1.01723	1.01559	1.00284	1.00113	0.00007	0.00008	10.0209	10.0083	0.0028	0.0039
NIST #3A	30	1.01598	1.01482	1.00161	1.00037	0.00005	0.00008	30.0316	30.0056	0.0065	0.0072
	100	1.01430	1.01416	0.999956	0.999717	0.000037	0.000069	99.9981	99.9781	0.0082	0.0176
(Hg UIM)	300			0.999860	0.999867	0.000021	0.000027	299.9700	299.9640	0.0147	0.0138
	1000			0.999926	0.999896	0.000005	0.000008	999.9337	999.9029	0.0097	0.0132

NMI	Target Press.	Calibration Ratios Before Re-scaling		a_{ij} or b_{ij}		u_{rdm} (a_{ij} or b_{ij})		p_{ij} / Pa		$u_c(p_{ij})$ / Pa	
				CDG1	CDG2	CDG1	CDG2	CDG1	CDG2	CDG1	CDG2
	Pa	CDG1	CDG2	RSG1	RSG2	RSG1	RSG2	RSG1	RSG2	RSG1	RSG2
	1	1.01451	1.01023	1.00224	1.00087	0.00038	0.00014	1.0026	1.0012	0.0013	0.0012
(Oil UIM)	3	1.01511	1.01108	1.00284	1.00171	0.00016	0.00011	3.0105	3.0060	0.0018	0.0017
	10	1.01397	1.01025	1.00170	1.00089	0.00008	0.00009	10.0245	10.0119	0.0025	0.0025
NIST #2B	30	1.01300	1.00951	1.00075	1.00016	0.00007	0.00009	30.0391	30.0100	0.0040	0.0056
	100	1.01223	1.00913	0.999987	0.999778	0.000065	0.000091	99.9961	99.9715	0.0101	0.0219
(Hg UIM)	300			0.999997	0.999945	0.000013	0.000021	299.9872	299.9797	0.0113	0.0093
	1000			0.999931	0.999899	0.000006	0.000004	999.9229	999.8929	0.0116	0.0190
	1	1.00789	1.01275	0.99733	0.99727	0.00067	0.00026	0.9977	0.9976	0.0013	0.0011
	3	1.01068	1.01501	1.00009	0.99950	0.00030	0.00025	3.0023	2.9993	0.0021	0.0021
	10	1.01012	1.01462	0.99953	0.99911	0.00017	0.00019	10.0028	9.9941	0.0062	0.0061
NPL-I	30	1.00997	1.01473	0.99939	0.99922	0.00022	0.00024	29.9984	29.9820	0.0180	0.0188
	100	1.00938	1.01452	0.998801	0.999014	0.000112	0.000142	99.8775	99.8951	0.054	0.060
	300			0.999185	0.999208	0.000064	0.000087	299.7435	299.7586	0.158	0.158
	1000			0.999016	0.999025	0.000178	0.000180	999.0078	999.0186	0.751	0.752
	1	1.22715	1.23163	1.21435	1.21285	0.02587	0.02551	1.2148	1.2132	0.0273	0.0271
	3	1.11467	1.11947	1.10304	1.10240	0.00955	0.00954	3.3114	3.3081	0.0311	0.0311
	10	1.06910	1.07414	1.05795	1.05776	0.00627	0.00628	10.5874	10.5808	0.0621	0.0622
CSIRO	30	1.04193	1.04713	1.03107	1.03117	0.00306	0.00307	30.9493	30.9406	0.0935	0.0937
	100	1.01889	1.02430	1.008272	1.008681	0.000702	0.000736	100.8246	100.8617	0.1042	0.1092
	300			1.001877	1.001945	0.000218	0.000229	300.5511	300.5796	0.1641	0.1651
	1000			1.000231	1.000238	0.000082	0.000076	1000.2228	1000.2313	0.1712	0.1701
	1	1.01570	1.01971	1.00378	1.0019	0.00024	0.00015	1.0042	1.0022	0.0013	0.0012
(Oil UIM)	3	1.01530	1.02009	1.00338	1.0022	0.00013	0.00009	3.0122	3.0075	0.0017	0.0017
	10	1.01335	1.01870	1.00145	1.0009	0.00008	0.00007	10.0220	10.0117	0.0025	0.0024
NIST #3B	30	1.01255	1.01823	1.00066	1.0004	0.00008	0.00007	30.0365	30.0174	0.0041	0.0053
	100	1.01203	1.01808	1.000144	1.000254	0.000066	0.000063	100.0118	100.0191	0.0102	0.0209
(Hg UIM)	300			0.999985	0.999913	0.000020	0.000032	299.9837	299.9699	0.0121	0.0118
	1000			0.999951	0.999932	0.000007	0.000012	999.9430	999.9259	0.0121	0.0221
	10	1.01142	1.01697	1.00190	1.00100	0.00049	0.00033	10.0265	10.0130	0.0141	0.0135
	30	1.01141	1.01732	1.00190	1.00134	0.00023	0.00019	30.0737	30.0457	0.0151	0.0154
KRISS	100	1.01010	1.01642	1.000595	1.000457	0.000131	0.000098	100.0570	100.0394	0.0214	0.0311
	300			1.000114	0.999966	0.000040	0.000049	300.0223	299.9859	0.0264	0.0257
	1000			0.999974	0.999942	0.000021	0.000021	999.9661	999.9358	0.0571	0.0607
	1	1.01268	1.01724	1.00384	1.00233	0.00051	0.00043	1.0042	1.0026	0.0014	0.0013
(Oil UIM)	3	1.01190	1.01693	1.00306	1.00202	0.00023	0.00018	3.0112	3.0069	0.0019	0.0019
	10	1.01013	1.01562	1.00131	1.00072	0.00010	0.00009	10.0206	10.0102	0.0029	0.0025
NIST #4B	30	1.00959	1.01535	1.00077	1.00046	0.00006	0.00007	30.0400	30.0191	0.0041	0.0062
	100	1.00897	1.01510	1.000161	1.000213	0.000046	0.000057	100.0135	100.0150	0.0109	0.0267
(Hg UIM)	300			1.000057	0.999924	0.000044	0.000054	300.0051	299.9733	0.0184	0.0179
	1000			0.999952	0.999945	0.000012	0.000017	999.9440	999.9384	0.0174	0.0299

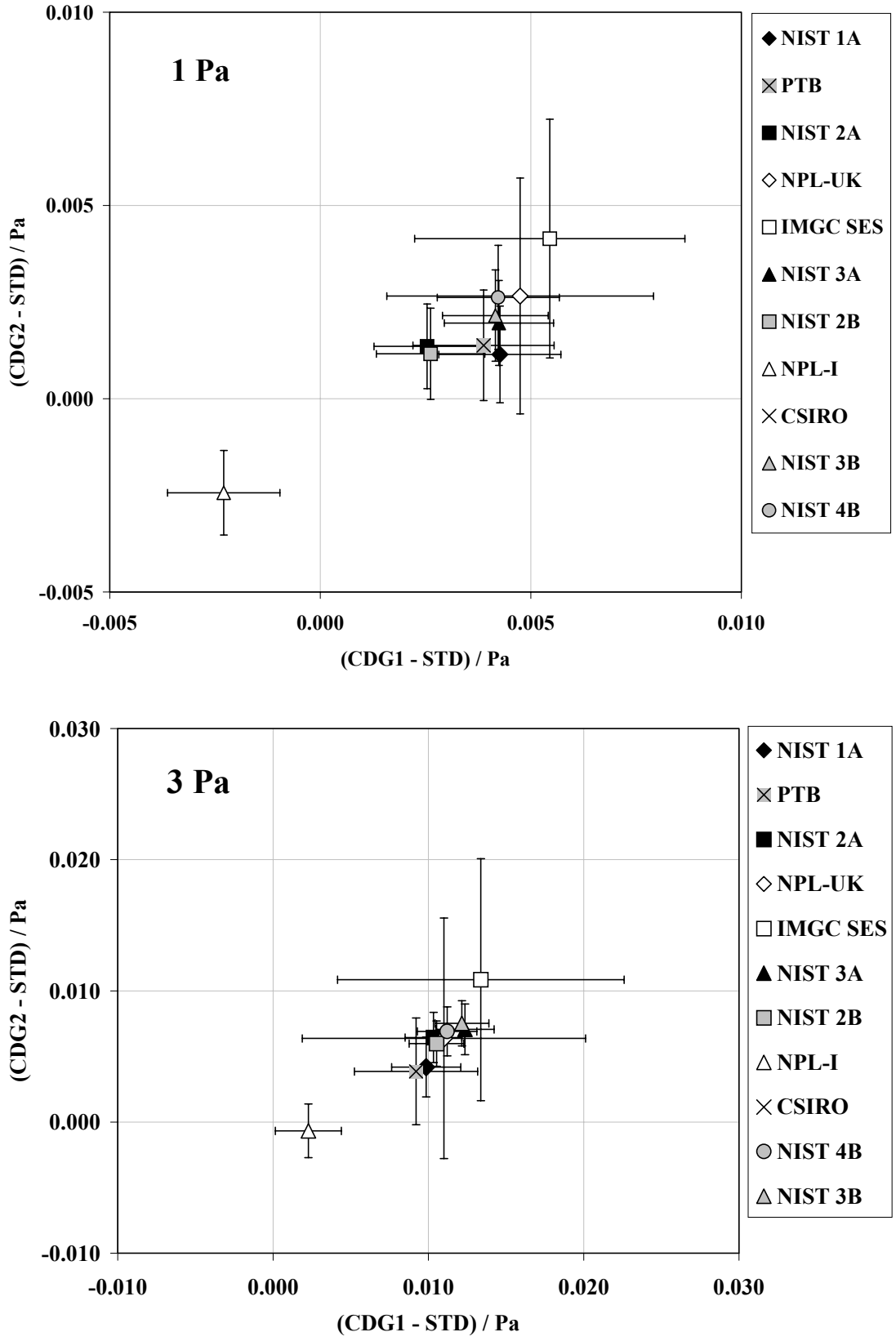


Figure 9. Youden plots of differences between normalized pressure readings of CDGs and pressures measured/generated by primary standards when equal to target pressures of 1 Pa and 3 Pa. The error bars refer to combined standard (k = 1) uncertainties.

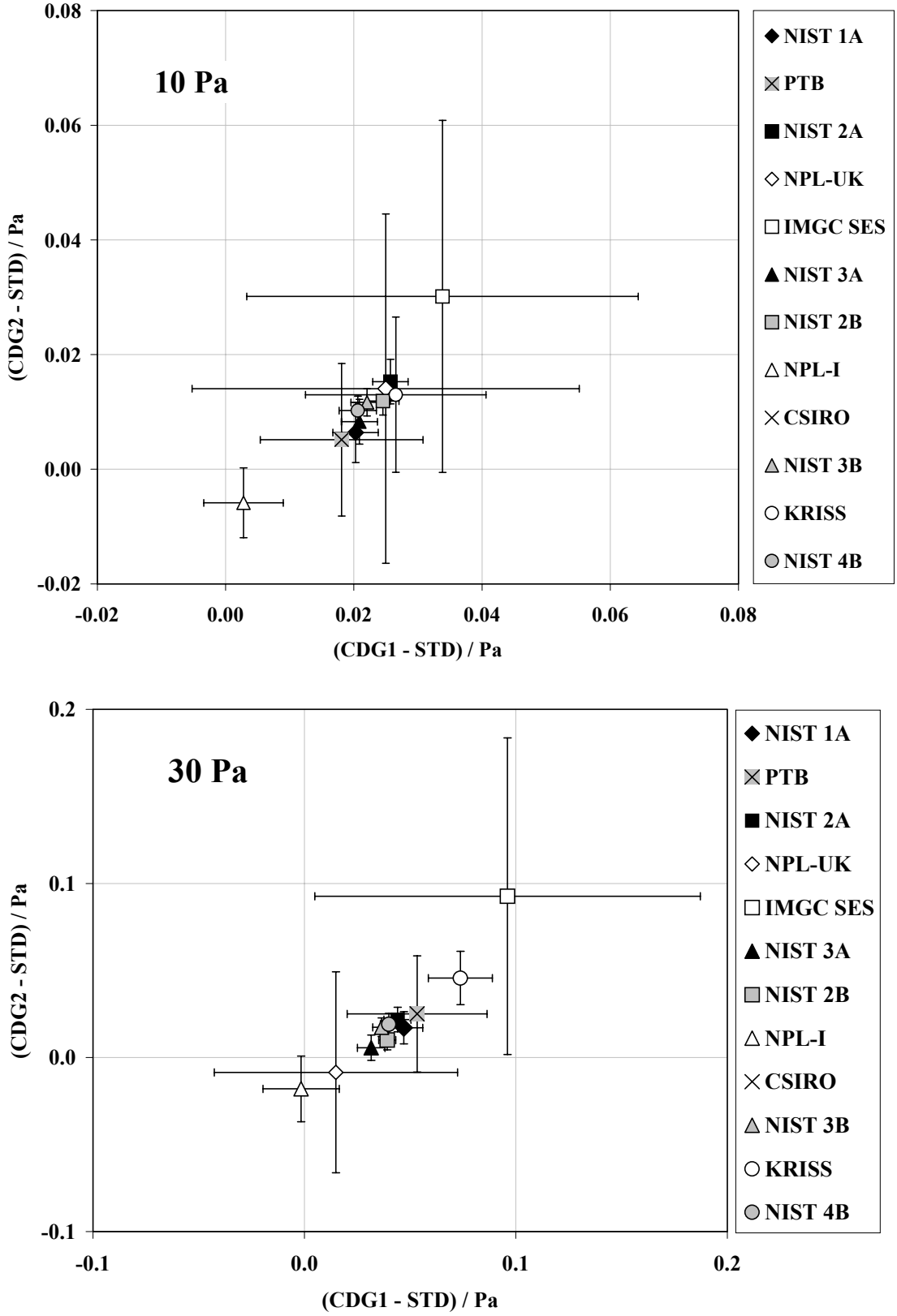


Figure 10. Youden plots of differences between normalized pressure readings of CDGs and pressures measured/generated by primary standards when equal to target pressures of 10 Pa and 30 Pa. The error bars refer to combined standard (k = 1) uncertainties.

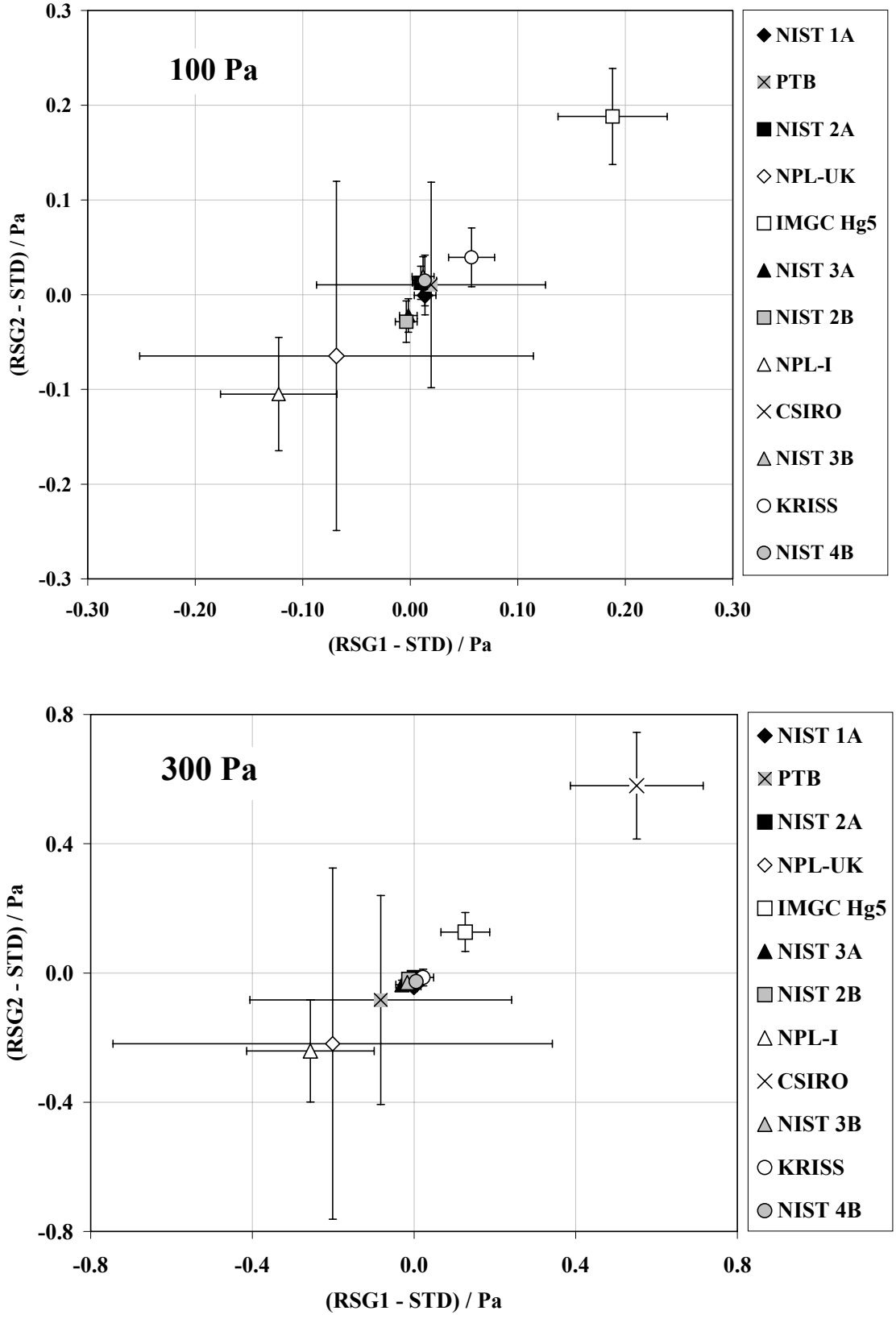


Figure 11. Youden plots of differences between normalized pressure readings of RSGs and pressures measured/generated by primary standards when equal to target pressures of 100 Pa and 300 Pa. The error bars refer to combined standard (k = 1) uncertainties.

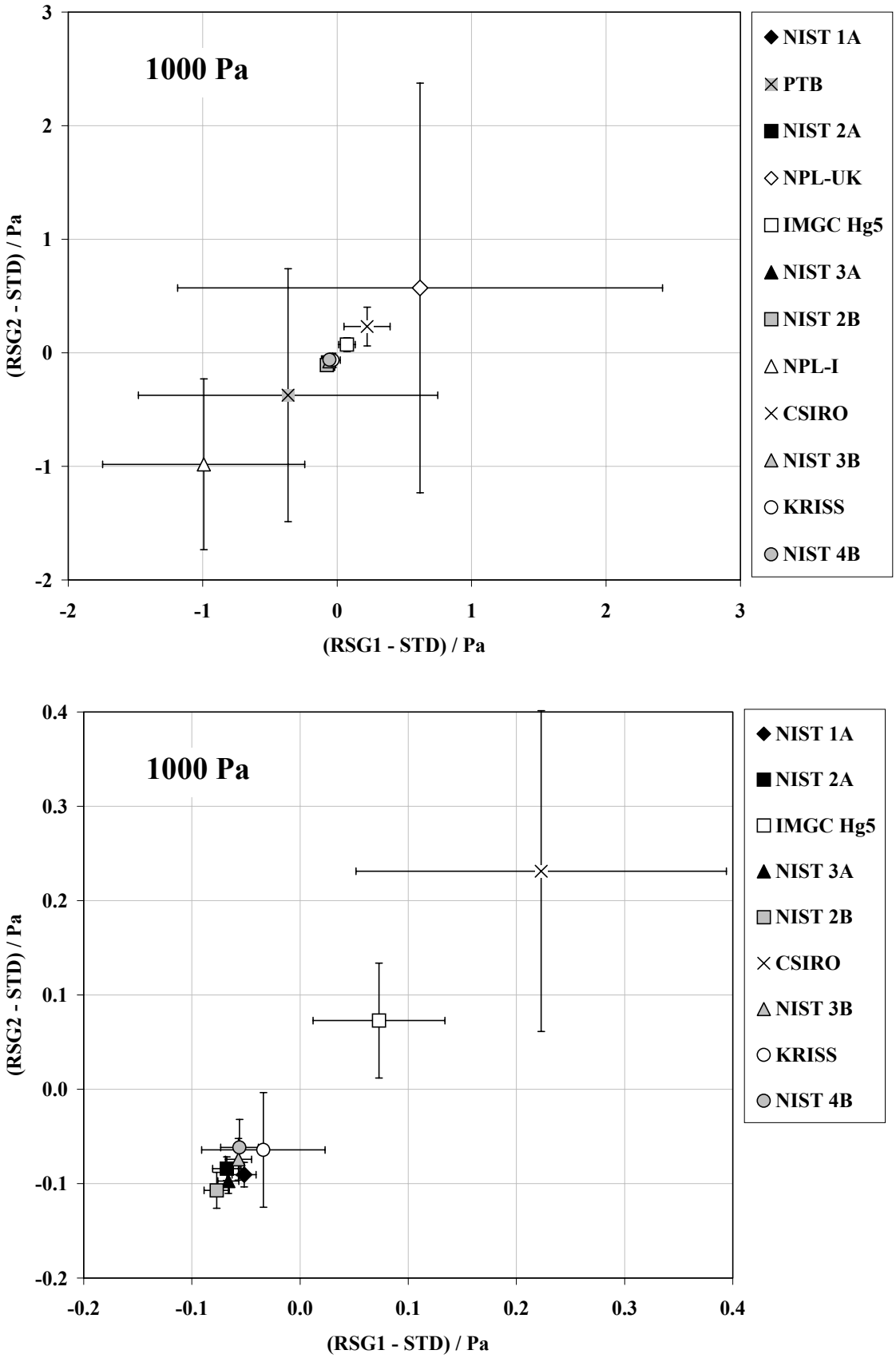


Figure 12. Youden plots of differences between normalized pressure readings of RSGs and pressures measured/generated by primary standards when equal to a target pressure of 1000 Pa. The error bars refer to combined standard ($k = 1$) uncertainties. The lower plot shows data for the liquid-column manometers only.

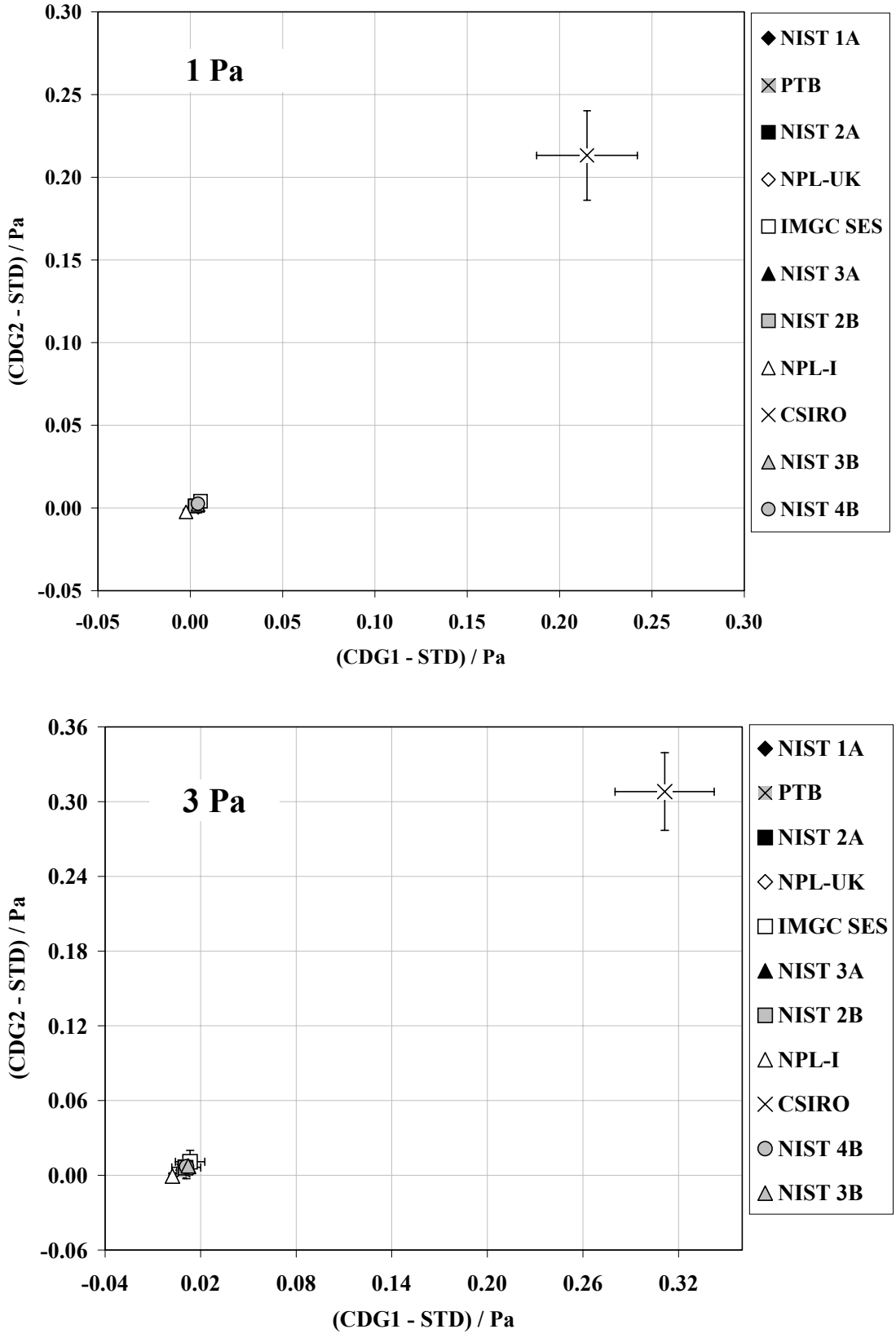


Figure 13. Youden plots from Figure 9 re-scaled to show data from all participating laboratories, including an outlier.

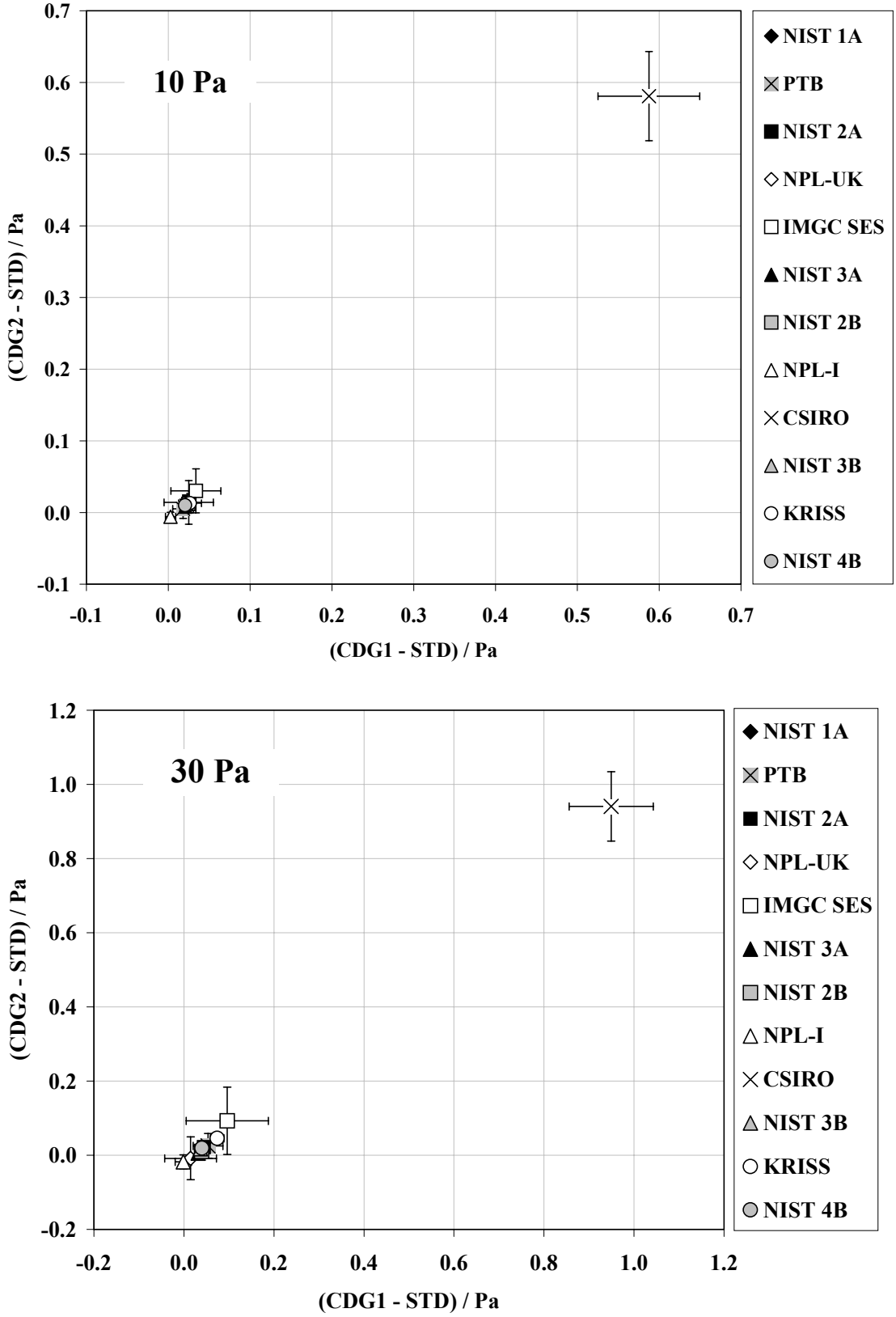


Figure 14. Youden plots from Figure 10 re-scaled to show data from all participating laboratories, including an outlier.

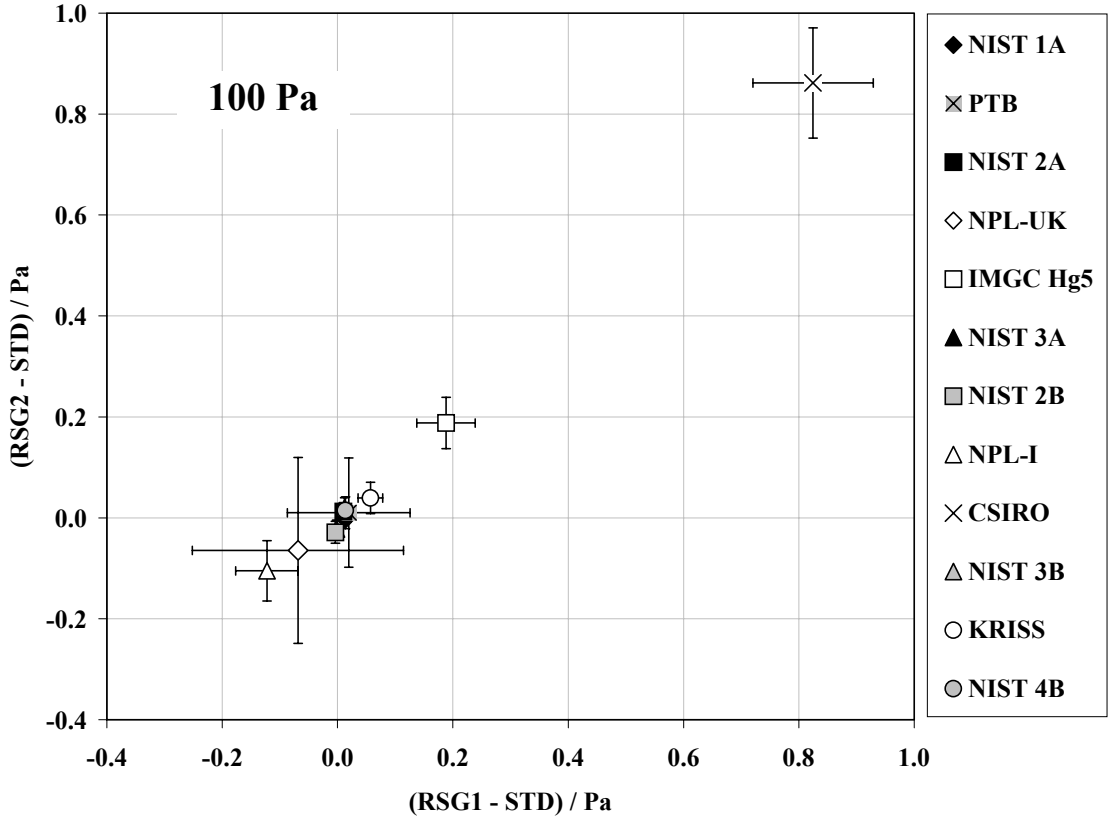


Figure 15. The upper Youden plot from Figure 11 re-scaled to show data from all participating laboratories, including an outlier.

7.2. DEGREES OF EQUIVALENCE OF THE PRIMARY STANDARDS

Table 6 presents a summary of final results for the pilot and participant NMIs as a function of nominal target pressures. The values for the corrected mean gauge readings p_j, which were calculated from Equation (17) using data in Tables 5 and A1, are given in column three. The combined uncertainties $u_c(p_j)$, which were calculated using Equation (26) or (27), are given in column four. The remaining columns present degrees of equivalence of the measurement standards expressed quantitatively in two ways: (1) deviations from reference values, and (2) pairwise differences between these deviations. The deviations D_j were calculated via Equations (28) and (A1) using data from Tables 6 and A1 in which the reference values do not include outliers seen in Figures 13, 14, and 15. The expanded uncertainties of these deviations, U_j, were calculated using Equation (29) and data in Tables 6 and A1, and the coverage factors in Table A2. The pairwise differences between the deviations, $D_{jj'}$, and the expanded uncertainties of these differences, $U_{jj'}$, were calculated using Equations (30) and (31) and data from Table 6, and coverage factors from Table A2. The shaded cells in Table 6 indicate pressures at which the condition $|D_j| \leq U_j$ or $|D_{jj'}| \leq U_{jj'}$ is not satisfied.

The degrees of equivalence of individual NMIs with respect to key comparison reference values are presented graphically in Figures 16 to 19 as plots of deviations, $D_j = p_j - p_R$, versus NMI and are summarized in Figure 20 where the ratios, D_j / U_j, for the participating laboratories are plotted as a function of pressure.

Table 6. Degrees of equivalence expressed in two ways: degrees of equivalence of NMIs relative to key comparison reference values, and degrees of equivalence between pairs of NMIs. D_j is the deviation of the corrected mean gauge reading p_j obtained by NMI_j from the reference value and U_j is the expanded uncertainty† of this deviation. $D_{jj'}$ is the difference between pairs of corrected mean gauge readings from NMI_j and $NMI_{j'}$, and $U_{jj'}$ is the expanded uncertainty† of this difference. The shaded cells indicate results for which $|D_j|$ exceeds U_j or $|D_{jj'}|$ exceeds $U_{jj'}$.

† U_j and $U_{jj'}$ refer to a 95% level of confidence

NMI_j	Nominal Pressure Pa	p_j Pa	$u_c(p_j)$ Pa	D_i Pa	U_i Pa	$NMI_{j'}$	CSIRO		IMGC		KRISS		NIST		NPL-I		NPL-UK		PTB	
							$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa
	1	1.211	0.023	0.211	0.046				0.209	0.046			0.211	0.046	0.216	0.046	0.210	0.046	0.211	0.046
	3	3.301	0.025	0.301	0.052				0.297	0.054			0.300	0.052	0.308	0.052	0.300	0.054	0.302	0.052
	10	10.566	0.046	0.57	0.10				0.55	0.11	0.56	0.10	0.57	0.10	0.58	0.10	0.56	0.11	0.57	0.10
CSIRO	30	30.906	0.069	0.91	0.15				0.85	0.23	0.88	0.15	0.92	0.15	0.95	0.15	0.94	0.18	0.90	0.16
	100	100.831	0.093	0.83	0.20				0.65	0.21	0.79	0.19	0.84	0.19	0.96	0.21	0.91	0.40	0.83	0.28
	300	300.57	0.16	0.57	0.34				0.44	0.33	0.56	0.31	0.58	0.31	0.81	0.44	0.8	1.1	0.65	0.70
	1000	1000.33	0.16	0.33	0.78				0.15	0.34	0.28	0.33	0.30	0.32	1.2	1.5	-0.4	3.5	0.6	2.2
	1	1.0023	0.0031	0.0023	0.0054		-0.209	0.046					0.0021	0.0064	0.0071	0.0063	0.0011	0.0085	0.0022	0.0067
(SES)	3	3.0042	0.0091	0.004	0.016		-0.297	0.054					0.003	0.018	0.011	0.018	0.003	0.025	0.006	0.019
	10	10.015	0.030	0.015	0.053		-0.55	0.11			0.012	0.065	0.015	0.060	0.033	0.061	0.012	0.084	0.020	0.064
IMGC	30	30.057	0.091	0.06	0.16		-0.85	0.23			0.03	0.18	0.07	0.18	0.10	0.18	0.09	0.21	0.06	0.19
	100	100.176	0.051	0.18	0.11		-0.65	0.21			0.14	0.13	0.18	0.11	0.30	0.15	0.25	0.37	0.17	0.23
(Hg-5)	300	300.132	0.060	0.13	0.24		-0.44	0.33			0.12	0.17	0.15	0.13	0.38	0.33	0.3	1.1	0.21	0.64
	1000	1000.178	0.061	0.18	0.74		-0.15	0.34			0.12	0.18	0.15	0.13	1.1	1.5	-0.5	3.5	0.4	2.2
	10	10.003	0.013	0.003	0.023		-0.56	0.10	-0.012	0.065			0.003	0.027	0.021	0.029	0.000	0.065	0.008	0.036
	30	30.022	0.014	0.022	0.035		-0.88	0.15	-0.03	0.18			0.032	0.028	0.069	0.044	0.06	0.12	0.020	0.070
KRISS	100	100.036	0.021	0.036	0.081		-0.79	0.19	-0.14	0.13			0.045	0.043	0.16	0.11	0.11	0.36	0.03	0.21
	300	300.010	0.023	0.01	0.22		-0.56	0.31	-0.12	0.17			0.023	0.047	0.25	0.31	0.2	1.1	0.09	0.63
	1000	1000.056	0.055	0.06	0.74		-0.28	0.33	-0.12	0.18			0.03	0.11	0.9	1.5	-0.6	3.5	0.3	2.2
	1	1.0003	0.0010	0.0003	0.0015		-0.211	0.046	-0.0021	0.0064					0.0051	0.0028	-0.0010	0.0063	0.0001	0.0035
(Oil UIM)	3	3.0008	0.0015	0.0008	0.0037		-0.300	0.052	-0.003	0.018					0.0079	0.0048	0.000	0.018	0.0022	0.0081
	10	9.9997	0.0018	0.000	0.013		-0.57	0.10	-0.015	0.060	-0.003	0.027			0.018	0.012	-0.003	0.059	0.005	0.025
NIST	30	29.9898	0.0026	-0.010	0.029		-0.92	0.15	-0.07	0.18	-0.032	0.028			0.037	0.035	0.02	0.11	-0.012	0.064
	100	99.9909	0.0053	-0.009	0.074		-0.84	0.19	-0.18	0.11	-0.045	0.043			0.12	0.11	0.07	0.36	-0.01	0.21
(Hg UIM)	300	299.9871	0.0054	-0.01	0.22		-0.58	0.31	-0.15	0.13	-0.023	0.047			0.23	0.31	0.2	1.1	0.06	0.63
	1000	1000.0308	0.0059	0.03	0.73		-0.30	0.32	-0.15	0.13	-0.03	0.11			0.9	1.5	-0.7	3.5	0.3	2.2

† U_j and $U_{jj'}$ refer to a 95% level of confidence

NMI_j	Nominal Pressure Pa	p_i Pa	$u_c(p_i)$ Pa	D_i Pa	U_i Pa	$NMI_{j'}$											
						CSIRO		IMGC		KRISS		NIST		NPL-I		NPL-UK		PTB	
						$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa	$D_{jj'}$ Pa	$U_{jj'}$ Pa
NPL-I	1	0.9952	0.0010	-0.0048	0.0023	-0.216	0.046	-0.0071	0.0063			-0.0051	0.0028			-0.0061	0.0063	-0.0050	0.0035
	3	2.9929	0.0019	-0.0071	0.0049	-0.308	0.052	-0.011	0.018			-0.0079	0.0048			-0.008	0.018	-0.0057	0.0084
	10	9.9817	0.0058	-0.018	0.016	-0.58	0.10	-0.033	0.061	-0.021	0.029	-0.018	0.012			-0.021	0.060	-0.013	0.027
	30	29.953	0.017	-0.047	0.041	-0.95	0.15	-0.10	0.18	-0.069	0.044	-0.037	0.035			-0.01	0.12	-0.049	0.072
	100	99.874	0.055	-0.13	0.11	-0.96	0.21	-0.30	0.15	-0.16	0.11	-0.12	0.11			-0.05	0.37	-0.13	0.23
	300	299.76	0.16	-0.24	0.33	-0.81	0.44	-0.38	0.33	-0.25	0.31	-0.23	0.31			0.0	1.1	-0.17	0.70
	1000	999.12	0.74	-0.9	1.4	-1.2	1.5	-1.1	1.5	-0.9	1.5	-0.9	1.5			-1.6	3.8	-0.6	2.6
NPL-UK	1	1.0013	0.0031	0.0013	0.0054	-0.210	0.046	-0.0011	0.0085			0.0010	0.0063	0.0061	0.0063			0.0011	0.0066
	3	3.0008	0.0091	0.001	0.016	-0.300	0.054	-0.003	0.025			0.000	0.018	0.008	0.018			0.002	0.019
	10	10.003	0.030	0.003	0.053	-0.56	0.11	-0.012	0.084	0.000	0.065	0.003	0.059	0.021	0.060			0.008	0.064
	30	29.966	0.057	-0.03	0.10	-0.94	0.18	-0.09	0.21	-0.06	0.12	-0.02	0.11	0.01	0.12			-0.04	0.13
	100	99.92	0.18	-0.08	0.30	-0.91	0.40	-0.25	0.37	-0.11	0.36	-0.07	0.36	0.05	0.37			-0.08	0.42
	300	299.80	0.54	-0.20	0.90	-0.8	1.1	-0.3	1.1	-0.2	1.1	-0.2	1.1	0.0	1.1			-0.1	1.2
	1000	1000.7	1.8	0.7	3.0	0.4	3.5	0.5	3.5	0.6	3.5	0.7	3.5	1.6	3.8			1.0	4.2
PTB	1	1.0002	0.0015	0.0002	0.0029	-0.211	0.046	-0.0022	0.0067			-0.0001	0.0035	0.0050	0.0035	-0.0011	0.0066		
	3	2.9987	0.0038	-0.0013	0.0075	-0.302	0.052	-0.006	0.019			-0.0022	0.0081	0.0057	0.0084	-0.002	0.019		
	10	9.995	0.013	-0.005	0.025	-0.57	0.10	-0.020	0.064	-0.008	0.036	-0.005	0.025	0.013	0.027	-0.008	0.064		
	30	30.002	0.033	0.002	0.062	-0.90	0.16	-0.06	0.19	-0.020	0.070	0.012	0.064	0.049	0.072	0.04	0.13		
	100	100.00	0.11	0.00	0.19	-0.83	0.28	-0.17	0.23	-0.03	0.21	0.01	0.21	0.13	0.23	0.08	0.42		
	300	299.92	0.32	-0.08	0.56	-0.65	0.70	-0.21	0.64	-0.09	0.63	-0.06	0.63	0.17	0.70	0.1	1.2		
	1000	999.7	1.1	-0.3	1.9	-0.6	2.2	-0.4	2.2	-0.3	2.2	-0.3	2.2	0.6	2.6	-1.0	4.2		

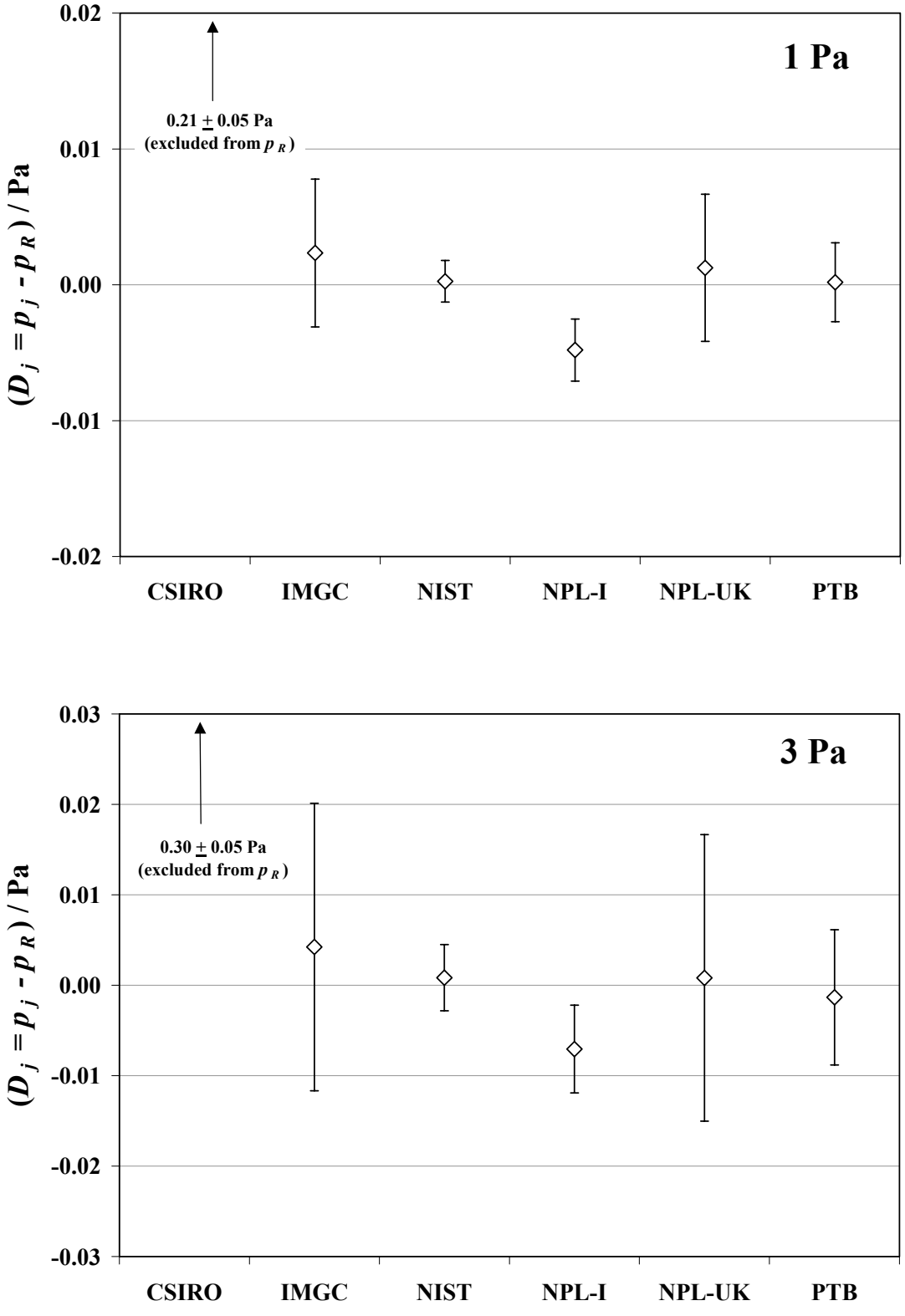


Figure 16. Degrees of equivalence expressed as deviations of corrected mean gauge readings from the key comparison reference values at 1 Pa and 3 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.

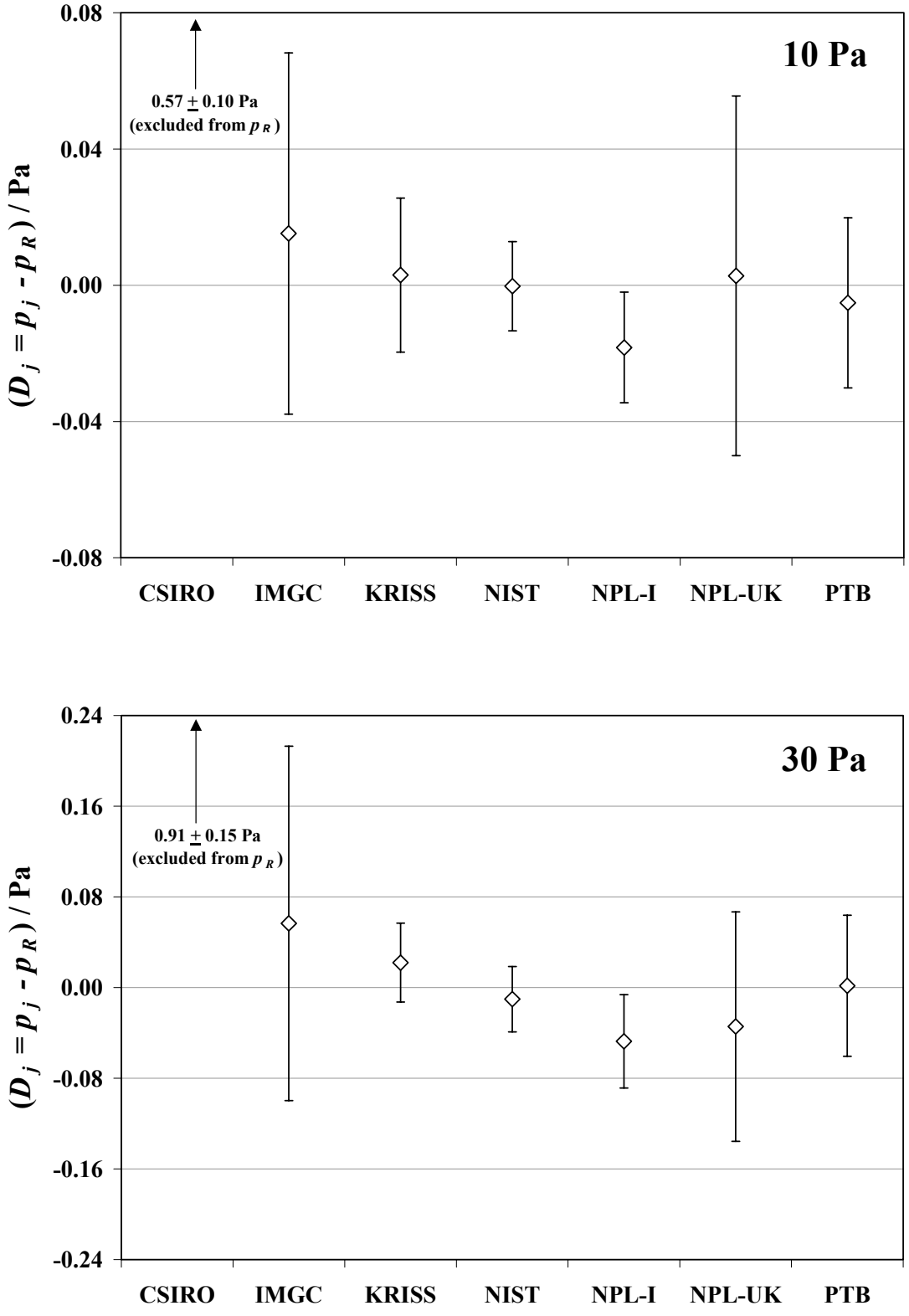


Figure 17. Degrees of equivalence expressed as the deviation of corrected mean gauge readings from the key comparison reference values at 10 Pa and 30 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.

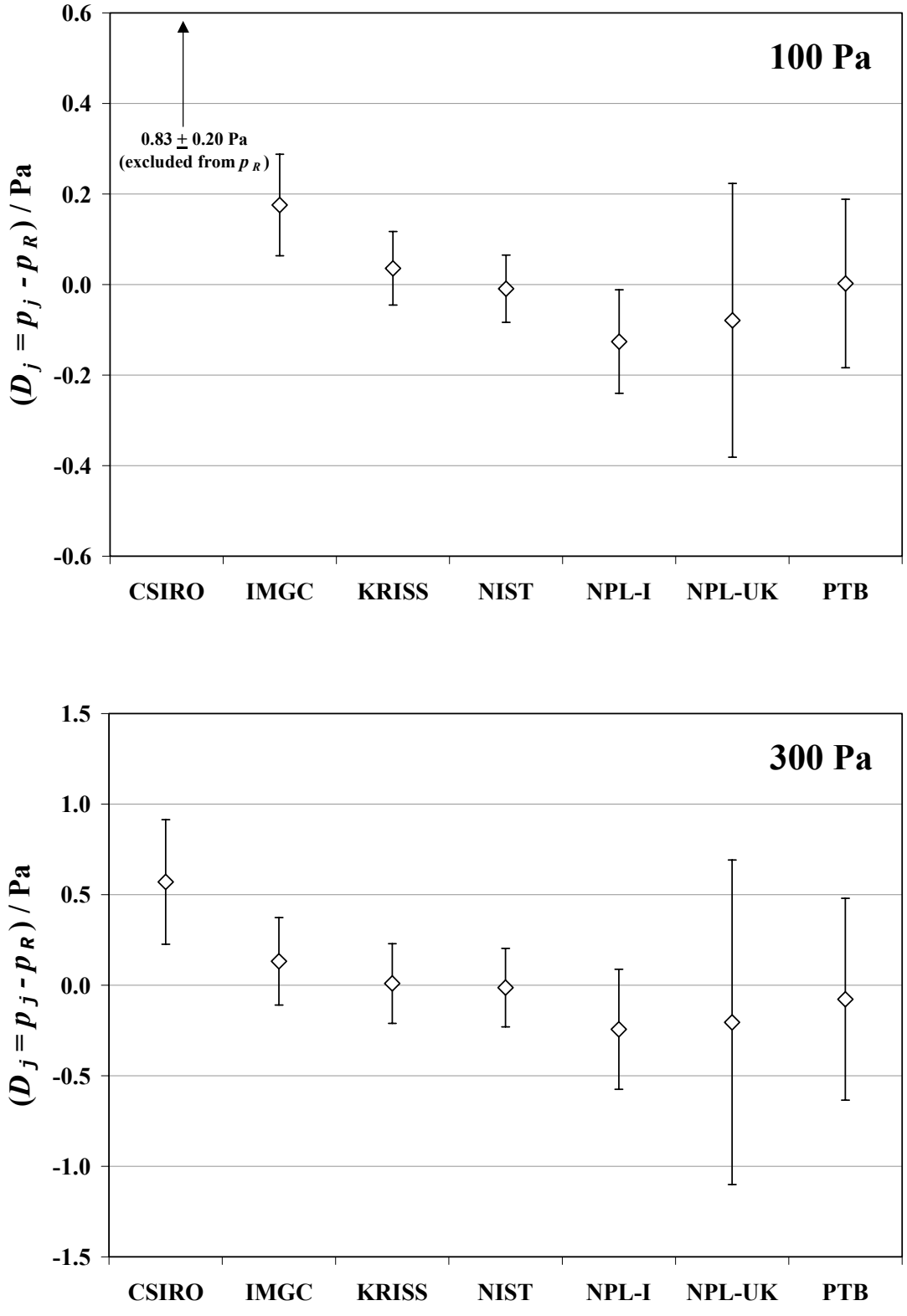


Figure 18. Degrees of equivalence expressed as the deviation of corrected mean gauge readings from the key comparison reference values at 100 Pa and 300 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.

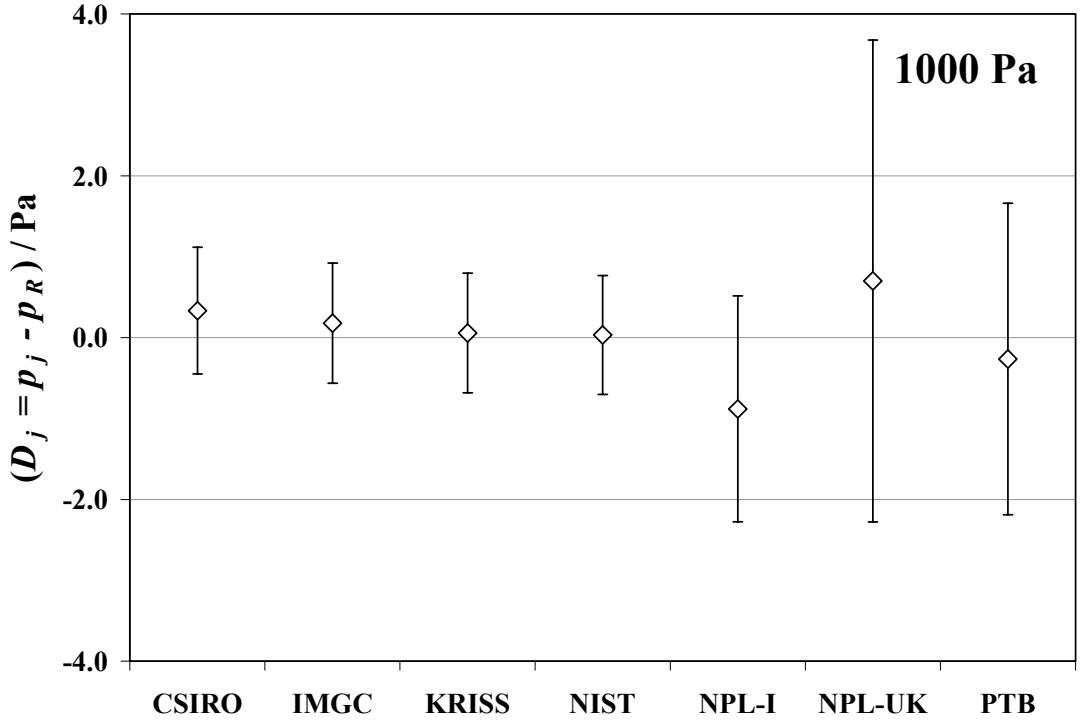


Figure 19. Degrees of equivalence expressed as the deviation of corrected mean gauge readings from the key comparison reference value at 1000 Pa. The error bars refer to expanded uncertainties of the deviations at a 95 % level of confidence.

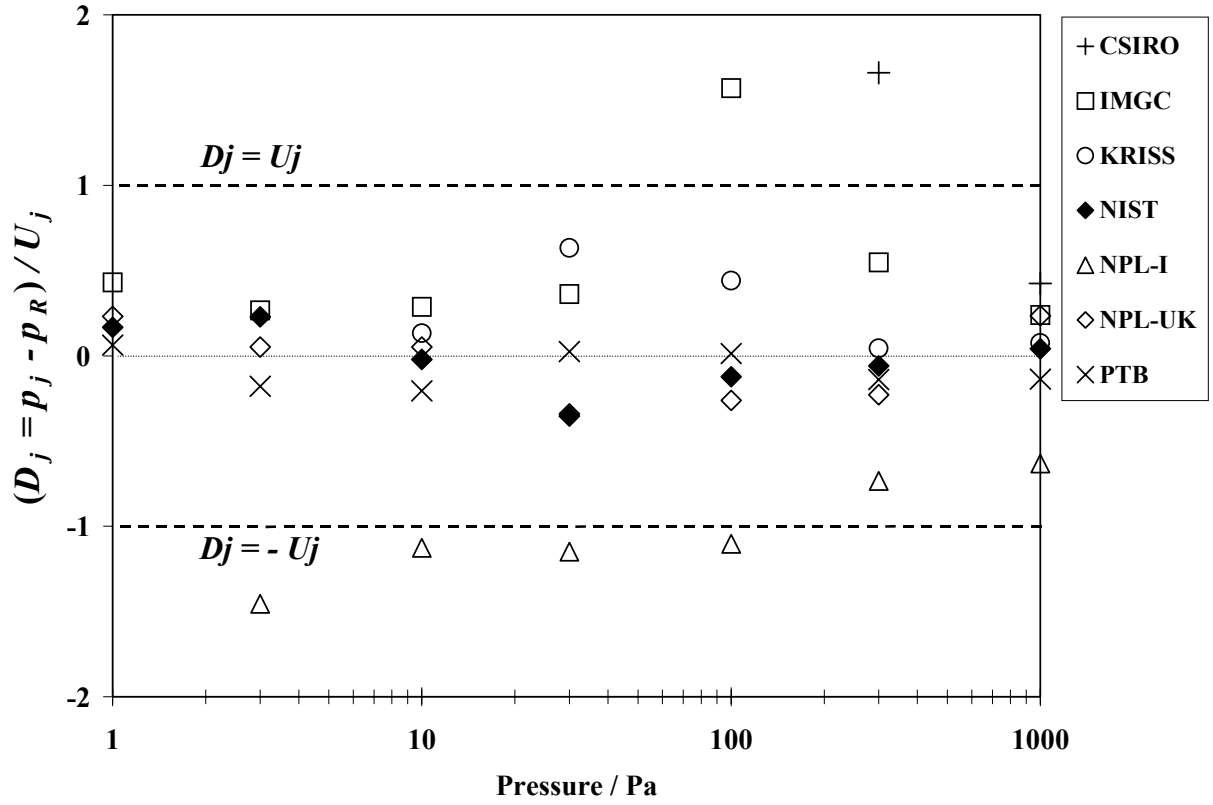


Figure 20. Summary of results for the degrees of equivalence of NMIs with respect to the key comparison reference values. The heavy dashed lines indicate the 95 % confidence interval.

8. DISCUSSION

The use of pairs of pressure transducers in the transfer standard package proved to be valuable not only because of redundancy (e.g., when RSG2 in Package A failed while at IMGC) but more importantly it enabled Youden graphical analyses to be used in the interpretation of the key comparison results. The Youden graphical representation has several important features. If only random errors of precision are present, the data points from individual primary standards will be distributed in a circular pattern (in the limit of a large number of standards). However if relative bias between individual primary standards exists, the data points will be distributed along a diagonal at 45 degrees to the positive *y*- and *x*-axes because primary standards that realize "true" pressures that are higher (or lower) will produce higher (or lower) readings in both pressure transducers. The scatter of data in a direction perpendicular to this diagonal provides a measure of precision of the transfer standard gauges. The Youden plots of the present results clearly show that the transfer standard gauges have sufficient precision to not only identify outliers (Figures 13 to 15) but also differentiate smaller relative systematic biases between individual primary standards (Figures 9 to 12).

In the present comparison, the degrees of equivalence of the measurement standards were expressed quantitatively in two ways: the deviation of corrected mean gauge readings from a key comparison reference value, D_j , and the pairwise difference between these deviations, $D_{jj'}$. When interpreting the results it is useful to note that $D_{jj'}$ may be regarded as a surrogate for the difference in "true" pressures actually realized by the pair of primary standards when both are set to measure/generate the same target pressure. Similarly, D_j represents the deviation of the "true" pressure realized by primary standard *j* from the corresponding key comparison reference value. However, D_j is not necessarily equal to the deviation of primary standard *j* from the SI value. Although the key comparison reference value is likely to be a close approximation to the SI value, it is possible that some of the results (corrected mean gauge readings) from individual NMIs may be even closer.

In the MRA the term 'degree of equivalence of measurement standards' is taken to mean the degree to which a standard is consistent with a key comparison reference value or with a measurement standard at another laboratory. A measure of the degree of consistency is provided by the relative magnitudes of the deviation D_j and its uncertainty, U_j , or the relative magnitudes of the pairwise difference $D_{jj'}$ and its uncertainty, $U_{jj'}$. The shaded cells in Table 6 indicate pressures at which results do not satisfy the condition $|D_j| \leq U_j$ or $|D_{jj'}| \leq U_{jj'}$. Clearly this is the case for results from the CSIRO at 1 Pa up to and including 100 Pa that were identified as outliers. However nearly half of the remaining cases that do not satisfy the condition are marginal and would satisfy the condition if results were rounded to one significant figure.

The pressures actually realized by the optical interferometer of the CSIRO at 300 Pa and below were significantly higher (~ 0.2 Pa to 0.9 Pa) than those realized by other primary standards (see Figures 13 to 15). Also, the pressures realized by the optical interferometer of the IMGC at 100 Pa and above were somewhat higher (~ 0.15 Pa) than pressures realized by the ultrasonic interferometer manometers (UIMs) at the KRISS and the NIST (see Figures 11 and 12). With the exception of the oil UIM at the NIST, the other manometers in this comparison used mercury as the manometric fluid[15].

The pressures realized by static expansion systems at the NPL-UK and the PTB were in good agreement with those realized at all other NMIs except CSIRO. The pressures realized by the expansion system at the NPL-I were generally in agreement with the other expansion systems though they tended to be systematically lower (~ 0.1% to 0.5%). However, the results at pressures of 100 Pa and below did not satisfy the condition $|D_j| \leq U_j$ with respect to the KCRVs nor the condition $|D_{jj'}| \leq U_{jj'}$ when compared with manometer primary standards, largely because of the relatively low uncertainty attributed to the NPL-I expansion system (see below).

---

[15] The unique properties of mercury, such as its relatively high vapor pressure (e.g., 0.17 Pa at 20 °C) and long diffusion times (many minutes to hours), need to be taken into account when determining the actual pressure being realized by a mercury manometer under absolute mode conditions. While the manometer generally does not measure its own vapor pressure, the transfer standard gauge will respond to the combined pressure of the local mercury vapor and the calibration gas.

The results presented in this report are based on data originally submitted to the pilot laboratory for preparation of the Draft A report and as such they represent the operational status of the low absolute-pressure standards at the time of the measurements[16]. After the circulation of Draft A, NPL-I re-examined their expansion system[17] and re-evaluated the uncertainties due to systematic effects. As a result the relative uncertainties were revised upward to (in parts per 10^2) 0.21 at 1 Pa, 0.18 at 3 Pa, 0.15 at 10 Pa and 30 Pa, and 0.11 at 100 Pa up to 1000 Pa. As may be seen in Table 3 and Figure 6, these uncertainties are comparable to those associated with expansion systems at the other NMIs. If the revised uncertainties are used instead of the earlier values, then the NPL-I results would satisfy the condition $|D_j| \leq U_j$ with respect to the KCRVs as well as the condition $|D_{jj'}| \leq U_{jj'}$ with respect to both other expansion systems and manometer primary standards, at essentially all pressures.

Also, CSIRO acknowledged that the corrections for mercury vapor pressure had not been included in their manometer data. If these corrections are applied, the CSIRO results would then also satisfy the condition $|D_j| \leq U_j$ or $|D_{jj'}| \leq U_{jj'}$ at pressures of 1 Pa and at 300 Pa and above. However significant disagreements would remain at pressures of 3 Pa, 10 Pa, 30 Pa, and 100 Pa.

9. CONCLUSIONS

The most critical element in the success of the present comparison of low absolute-pressure standards was the use of two different types of transducer as transfer standard artifacts. The combination of resonant silicon gauges with their exceptional calibration stability and capacitance diaphragm gauges to provide pressure resolution yielded transfer standards that had accuracies commensurate with the measurement standards being compared, over the entire pressure range of the comparison. In addition, the gauges were sufficiently rugged to withstand the inevitable rough treatment during shipment between laboratories.

The comparison tested two principal methods used by NMIs to realize their low absolute-pressure standards, namely, static expansion systems and liquid-column manometers. The results for four expansion systems and four manometers revealed no significant relative bias between the principal measurement methods.

Finally, the key comparison established the degrees of equivalence of absolute-pressure standards at seven NMIs, both with respect to key comparison reference values as well as between pairs of measurement standards. Apart from results from one NMI that were identified as outliers, the absolute-pressure standards of the participating NMIs were generally found to be equivalent.

---

[16] The Guidelines for CIPM Key Comparisons (Appendix F of the MRA) state that once results have been submitted to the pilot laboratory they stand and can only be changed under unusual circumstances and with the agreement of all participants.

[17] NPL-I identified several factors contributing to their results being systematically lower. In particular, the NPL-I UIM was not in operation at the time of the comparison and so the quartz spiral Bourdon gauge (QSBG) used for measuring the initial pressure as well as the 10 torr CDG used for measuring volume ratios had not been calibrated against the UIM for some time. Several months after the comparison data were submitted to the pilot laboratory, the UIM was made operational once again and the QSBG and CDG were recalibrated. Differences of ranging from 0.03 to 0.12 parts in 10^2 were found in the calibration factor of the QSBG. Also an improved value for the 25-mL volume was obtained.

APPENDIX

A1. REFERENCE VALUES FOR KEY COMPARISON CCM.P-K4

A key comparison reference value (KCRV) may be defined at each target pressure as an average of the mean normalized gauge readings that would be obtained at the different laboratories when their primary standards measure/generate pressures exactly equal to the target value. There are several procedures possible for averaging [28], which include a simple mean of all data, a median of all data and, since the primary standards represent two principal measurement methods, a mean of the measurement method means, or a weighted mean of the measurement method means with weights inversely proportional to estimates of method variance. Each procedure has some advantages and disadvantages.

An arithmetic mean of the combined data has the advantage of simplicity but if the "true" means of the different measurement methods are not the same but have relative bias then the arithmetic mean will weight the methods by "popularity", which is not desirable. Another disadvantage is the simple mean is sensitive to outliers.

A median of all data is relatively insensitive to outliers but it may effectively omit one of the measurement methods from the analysis if there is significant relative bias between methods. The major disadvantage of a median however is the lack of theory on which to base uncertainty estimates.

A major advantage of a mean of the measurement method means is that it incorporates the range of typical values obtained with different measurement methods without weighting by popularity, as does the simple mean of all data. However like the mean of the combined data, the mean of the method means is sensitive to the influence of outliers, which can only be eliminated by exclusion from the calculation of the method means.

A weighted mean of the measurement method means, with weights inversely proportional to the "true" method variance, may yield the most precise estimate of the overall mean but the weights must be known without error and any between-method bias must be negligible. Since weights are usually not known without error, using a weighted mean when weights are not known can lead to estimates with less precision than methods based on equal weights. Furthermore, weights (method variance) cannot be estimated when a method is represented by only one primary standard, as is the case at several pressures in the present comparison.

Considering these options, an unweighted mean of the measurement method means was selected as a reasonable procedure to obtain reference values for this key comparison.

As stated earlier (Section 6.5), the "true" pressures realized by the primary standards when set to measure/generate a given target pressure should, on average, closely approximate the SI value under the assumption that deviations from the SI value are randomly distributed. Therefore, it is reasonable to correct the mean normalized gauge readings so that their ensemble average (i.e., the KCRV) is also equal to the target pressure. This correction, in effect, sets the KCRV numerically equal to the target pressure.

At target pressures (p_t) of 1000 Pa and lower, the key comparison reference value p_R may be expressed in terms of the mean of measurement method means of corrected mean gauge readings ($p_j = f_C p_{jU}$) as follows:

$$p_R = \frac{1}{2}\left(\frac{1}{N_{MAN}} \sum_{j=1}^{N_{MAN}} p_j + \frac{1}{N_{SES}} \sum_{j=1}^{N_{SES}} p_j \right) = \frac{f_C}{2}\left(p_{MAN} + p_{SES} \right) \qquad \text{(A1)}$$

where f_C is the required correction factor such that $p_R = p_t$, and the measurement method means of the uncorrected gauge readings for N_{MAN} liquid-column manometers, p_{MAN}, and for N_{SES} static expansion systems, p_{SES}, are calculated from

$$p_{MAN} = \frac{1}{N_{MAN}} \sum_{j=1}^{N_{MAN}} p_{jU} \quad \text{and} \quad p_{SES} = \frac{1}{N_{SES}} \sum_{j=1}^{N_{SES}} p_{jU} \qquad \text{(A2)}$$

The uncertainties in the method means could be estimated from their sample variances

Table A1. Correction factors, f_C , reference values, p_R , and their estimated combined standard uncertainties, $u_c(p_R)$, calculated when excluding outliers at pressures up to and including 100 Pa (results above the dotted line)†. The measurement method means of data obtained with static expansion systems, p_{SES}, and with liquid-column manometers, p_{MAN} , and estimates of their respective standard uncertainties, u_{SES} and u_{MAN} , are also given. Not all digits are significant but are retained for calculation of final results.

Target Pressure Pa	N_{SES}	p_{SES} **Pa**	u_{SES} **Pa**	N_{MAN}	p_{MAN} **Pa**	u_{MAN} **Pa**	f_C	p_R **Pa**	$u_c(p_R)$ **Pa**
1	4	1.0022	0.0012	1	1.0027	0.0010	0.99756	1	0.0008
3	4	3.0070	0.0034	1	3.0087	0.0015	0.99738	3	0.0019
10	4	10.0154	0.0112	2	10.0181	0.0068	0.99833	10	0.0066
30	4	30.0317	0.0283	2	30.0436	0.0072	0.99875	30	0.015
100	3	99.9448	0.0730	3	100.0798	0.0184	0.99988	100	0.038
300	3	299.8194	0.2167	4	300.169	0.043	1.00002	300	0.11
1000	3	999.746	0.748	4	1000.044	0.045	1.000105	1000	0.37

† Results from CSIRO at pressures of 1 Pa up to and including 100 Pa appear to be outliers.

$$u_{MAN}^2 = \sum_{j=1}^{N_{MAN}} \frac{(p_{jU} - p_{MAN})^2}{N_{MAN}(N_{MAN} - 1)} \quad \text{and} \quad u_{SES}^2 = \sum_{j=1}^{N_{SES}} \frac{(p_{jU} - p_{SES})^2}{N_{SES}(N_{SES} - 1)} \tag{A3}$$

if the "true" means of data taken with individual primary standards of a given method were equal without between-standard bias. This condition would be satisfied if primary standards of a given method were constructed to be exact replicates. However, primary standards at the participant laboratories were developed "in-house" and as such, each standard of a given method is unique even though they share a common basic operating principle. The Youden plots in Section 7.1 indicate that there is some relative bias between individual primary standards of a given method, the outliers notwithstanding. Equation (A3) would also be applicable if the number of standards of a given method were sufficiently large so that any between-standard biases could be regarded as being randomly distributed. But in this comparison, the sample size is too limited (one to four depending on method and target pressure) and clearly Equation (A3) has no meaning for a sample size of one.

Alternatively, the uncertainties in p_{MAN} and p_{SES} may be regarded as arising from the propagation of uncertainties $u_c(p_{jU})$ associated with independent values of p_{jU} in which case

$$u_{MAN}^2 = \sum_{j=1}^{N_{MAN}} \frac{u_c^2(p_{jU})}{N_{MAN}^2} \quad \text{and} \quad u_{SES}^2 = \sum_{j=1}^{N_{SES}} \frac{u_c^2(p_{jU})}{N_{SES}^2} \tag{A4}$$

Table A1 presents values for p_{MAN} and p_{SES} and their uncertainties, which were calculated from Equations (A2), (16), and (A4) using data in Table 5 for the case when outliers at pressures of 1 Pa, 3 Pa, 10 Pa, 30 Pa and 100 Pa are excluded (see Figures 13, 14, and 15).

Using the approximation $f_C \simeq 1$ (see Table A1), the combined uncertainty in p_R can then be estimated from:

$$u_c^2(p_R) = (f_c/2)^2 (u_{MAN}^2 + u_{SES}^2) \simeq (1/2)^2 (u_{MAN}^2 + u_{SES}^2) \tag{A5}$$

where the uncertainties of the method means are given by Equation (A4) and the correction factor that sets the reference value equal to the target pressure is defined by Equation (A1) as

$$f_C \equiv 2p_t / (p_{MAN} + p_{SES}) \tag{A6}$$

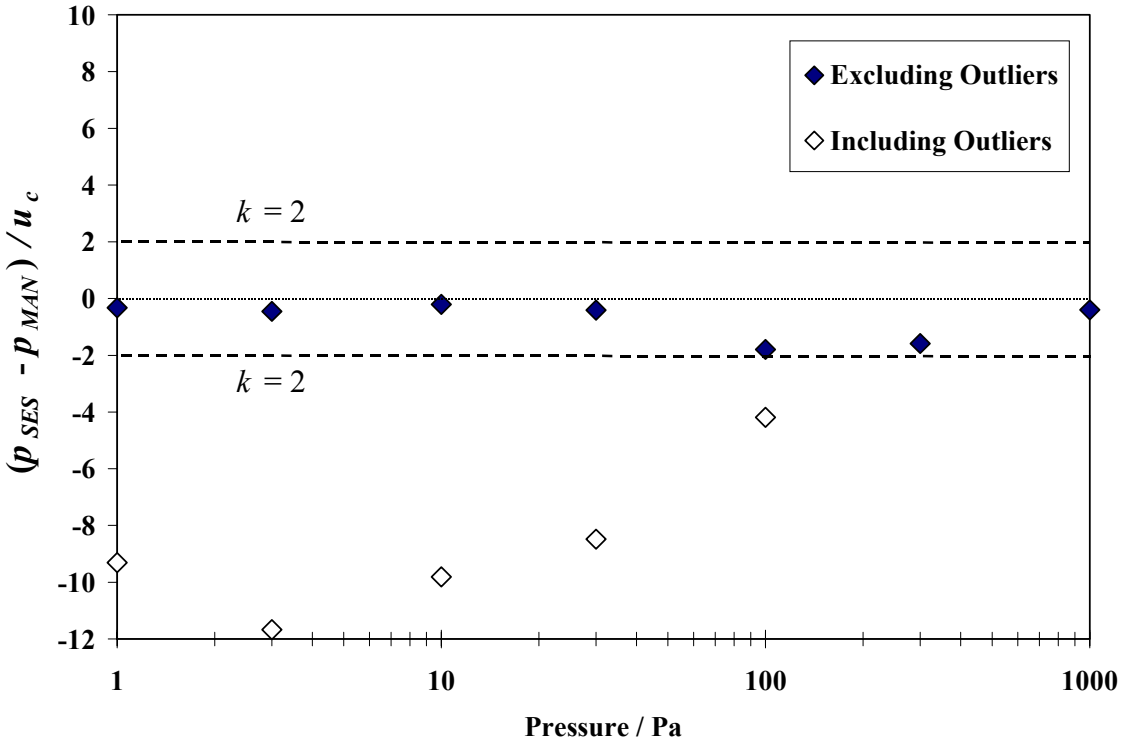


Figure A1. The ratio of the difference between measurement method means, $p_{SES} - p_{MAN}$, and the standard uncertainty of this difference as a function of pressure.

Implicit in Equation (A5) is the assumption that any between-method bias is small and can be neglected. Between-method bias can arise when there are unknown systematic effects (biases) in one or both measurement methods that have not been taken into account in the stated uncertainties for primary standards of a given method. If the between-method bias is significant, its contribution to the reference value (i.e., the mean of the method means) can be estimated by a Type B evaluation (e.g., [29]) and added in quadrature to uncertainties of the method means in Equation (A5).

In order to check this assumption it is instructive to compare the difference between measurement method means, $p_{SES} - p_{MAN}$, with its combined uncertainty, which can be estimated from

$$u_c^2 = u_{MAN}^2 + u_{SES}^2 \qquad \text{(A8)}$$

Figure A1 presents the ratio, $(p_{SES} - p_{MAN}) / u_c$, as a function of pressure for two cases, when outliers are excluded and when they are included in the calculation of p_{MAN} and u_{MAN}. The plot shows that, when outliers are excluded, the difference between measurement method means lies within two times its combined standard uncertainty at all pressures and, when they are included, the difference between methods significantly exceeds the $k = 2$ level at the associated pressures. The plot clearly shows the relative bias between the two methods to be relatively small and for this reason its contribution to uncertainty in the reference value was neglected.

Table A1 also presents calculated values for the correction factors, the reference values and their combined uncertainties for the case in which outliers (results from CSIRO at target pressures of 1 Pa up to and including 100 Pa) have been excluded.

A2. CALCULATION OF EFFECTIVE DEGREES OF FREEDOM AND ASSOCIATED COVERAGE FACTORS

In the present comparison, a conventional procedure described in references [25, 26] was used to calculate coverage factors k_{95} that provide uncertainty intervals $U_j = k_{95}\,u_c(D_j)$ and $U_{jj'} = k_{95}\,u_c(D_{jj'})$, as defined in Equations (29) and (31), with a level of confidence approximating 95 %. The first step was to estimate the "effective degrees of freedom" ν_{eff} for the combined standard uncertainties $u_c(D_j)$ and $u_c(D_{jj'})$. Values of ν_{eff} were estimated by combining the degrees of freedom of individual component uncertainties using the Welch-Satterthwaite formula. The second step was to obtain the *t*-factor, $t_{95}(\nu_{eff})$, by interpolating a table of values given in the above cited references and then take $k_{95} = t_{95}(\nu_{eff})$.

The degrees of freedom of a component uncertainty obtained from a Type A evaluation can be readily determined by appropriate statistical methods. In the case of the uncertainties due to short-term random effects [see Equations (19) and (22)] the number of degrees of freedom is $\nu_{rdm} = 5 - 1 = 4$.

The degrees of freedom to associate with a component uncertainty obtained from a Type B evaluation is more problematic. However, if the component uncertainties are chosen in such a way that the probability of the measurand lying outside these limits is extremely small (e.g., when uncertainties are obtained from a rectangular probability distribution), then the degrees of freedom become infinitely large. This approximation, which cannot be fully justified for the estimates of $u_{std}(p_{ij})$ and $u_{lts}(p_{ij})$ in this comparison, is not necessarily unrealistic since the Type B evaluations were generally carried out in a manner that attempted to avoid an underestimation of the component uncertainties.

In the approximation that $\nu_{std} \to \infty$ and $\nu_{lts} \to \infty$, the effective degrees of freedom of the combined standard uncertainty $u_c(D_j)$ associated with the deviation of a primary standard from the KCRV was estimated from the Welch-Satterthwaite formula as:

$$\nu_{eff} = \frac{u_c^4(D_j)\nu_{rdm}}{\sum_{i=1}^{2}\left(1-\frac{1}{N}\right)^2 c^4 u_{rdm}^4(p_{ij}) + \sum_{j'=1}^{N_{MAN}}\sum_{i=1}^{2}\frac{c^4}{16N_{MAN}^4}u_{rdm}^4(p_{ij'}) + \sum_{j'=1}^{N_{SES}}\sum_{i=1}^{2}\frac{c^4}{16N_{SES}^4}u_{rdm}^4(p_{ij'})} \qquad \text{(A8)}$$

where N is either N_{MAN} or N_{SES} depending on whether primary standard j is a manometer or an expansion system, and c is the (common) value for partial derivatives as defined in Equations (26) and (27). When the primary standard index j or j' refers to the pilot laboratory, the appropriate terms in the denominator must also include summations over the calibration number n and package label m as in Equation (27).

Similarly, the effective degrees of freedom of the combined standard uncertainty $u_c(D_{jj'})$ associated with the difference between primary standards was estimated from:

$$\nu_{eff} = \frac{u_c^4(D_{jj'})\nu_{rdm}}{\sum_{i=1}^{2} c^4 u_{rdm}^4(p_{ij}) + \sum_{i=1}^{2} c^4 u_{rdm}^4(p_{ij'})} \qquad \text{(A9)}$$

where conditions affecting the summation of terms in the denominator of Equation (A8) apply equally well to Equation (A9).

Estimated values for the effective degrees of freedom and associated coverage factors k_{95}, which are needed to calculate expanded uncertainties $U_j = k_{95}\,u_c(D_j)$ and $U_{jj'} = k_{95}\,u_c(D_{jj'})$ in Table 6 of Section 7.2, are presented in Table A2.

Table A2. Estimates of the "effective degrees of freedom" ν_{eff} associated with standard uncertainties, $u_c(D_j)$ and $u_c(D_{jj'})$, and the coverage factors k_{95} that produce expanded uncertainties, $U_j = k_{95}\,u_c(D_j)$ and $U_{jj'} = k_{95}\,u_c(D_{jj'})$, with an approximate 95 % level of confidence.

				NMI$_{j'}$											
NMI$_j$	**Nominal Pressure**	U_j		**IMGC** $U_{jj'}$		**KRISS** $U_{jj'}$		**NIST** $U_{jj'}$		**NPL-I** $U_{jj'}$		**NPL-UK** $U_{jj'}$		**PTB** $U_{jj'}$	
	Pa	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}
	1	42	2.02	44	2.01			42	2.02	42	2.02	44	2.01	42	2.02
	3	28	2.05	36	2.03			28	2.05	28	2.05	36	2.03	29	2.04
	10	12	2.18	24	2.06	14	2.14	12	2.18	12	2.18	24	2.06	13	2.16
CSIRO	30	13	2.16	85	1.99	13	2.16	12	2.18	13	2.16	33	2.03	17	2.11
	100	125	1.98	113	1.98	102	1.98	93	1.99	166	1.98	2196	1.96	495	1.97
	300	2680	1.96	520	1.97	1021	1.96	982	1.96	3839	1.96	160795	1.96	16947	1.96
	1000	35614	1.96	379	1.97	677	1.97	545	1.97	9911	1.96	8448819	1.96	534714	1.96
	1	25994	1.96					27317	1.96	14853	1.96	87327	1.96	31845	1.96
(SES)	3	69978	1.96					66484	1.96	59661	1.96	248544	1.96	81734	1.96
	10	70486	1.96			49961	1.96	63249	1.96	65834	1.96	246318	1.96	84157	1.96
IMGC	30	180249	1.96			156843	1.96	168864	1.96	154574	1.96	323549	1.96	210789	1.96
	100	71	2.00			6	2.45	22	2.08	17947	1.96	4232	1.96	632	1.96
(Hg-5)	300	427	1.98			5	2.57	14	2.14	30239	1.96	95437	1.96	9002	1.96
	1000	32502	1.96			13	2.16	13	2.16	16653	1.96	9473432	1.96	544043	1.96
	10	6493	1.96					3129	1.96	4137	1.96	109345	1.96	10363	1.96
	30	7540	1.96					850	1.97	2144	1.96	203638	1.96	26902	1.96
	100	5783	1.96					379	1.97	7905	1.96	1263056	1.96	174313	1.96
KRISS	300	126214	1.96					298	1.97	60098	1.96	8897513	1.96	31870	1.96
	1000	673962	1.96					1560	1.96	9490	1.96	209643107	1.96	873707	1.96
	3000	442	1.98					241	1.97						
	10000	335	1.99					241	1.97						

NMI_i	Nominal Pressure Pa	U_i		$NMI_{j'}$									
				IMGC		KRISS		NIST		NPL-I		NPL-UK		PTB	
				$U_{jj'}$		$U_{jj'}$		$U_{jj'}$		$U_{jj'}$		$U_{jj'}$		$U_{jj'}$	
		ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}	ν_{eff}	k_{95}
	1	182477	1.96							649	1.96	12721733	1.96	67629	1.96
(Oil UIM)	3	359914	1.96							1107	1.96	17351384	1.96	40438	1.96
	10	43862	1.96							1969	1.96	5738689	1.96	72173	1.96
NIST	30	147670	1.96							708	1.96	1324028	1.96	145204	1.96
	100	8593	1.96							5139	1.96	3679290	1.96	670162	1.96
(Hg UIM)	300	228566	1.96							32119	1.96	9955848	1.96	31660	1.96
	1000	737498	1.96							4694	1.96	238435428	1.96	872500	1.96
	1	994	1.96									32494	1.96	2801	1.96
	3	4184	1.96									454186	1.96	14475	1.96
	10	19343	1.96									1795159	1.96	52056	1.96
NPL-I	30	5006	1.96									165711	1.96	24166	1.96
	100	18365	1.96									1127368	1.96	192370	1.96
	300	168395	1.96									5607271	1.96	47282	1.96
	1000	17933	1.96									448870	1.96	94063	1.96
	1	11747248	1.96											869869	1.96
	3	16486634	1.96											1232866	1.96
	10	4142737	1.96											2243716	1.96
NPL-UK	30	1457076	1.96											1177881	1.96
	100	1507097	1.96											4036121	1.96
	300	9615486	1.96											461242	1.96
	1000	114681004	1.96											11214095	1.96
	1	57016	1.96												
	3	51338	1.96												
	10	105223	1.96												
PTB	30	215406	1.96												
	100	249026	1.96												
	300	43698	1.96												
	1000	1155404	1.96												

ACKNOWLEDGEMENTS

Contributions by several members of the Pressure and Vacuum Group at the NIST are gratefully acknowledged, in particular, S. A. Tison and C. R. Tilford, for their help with development of the transfer standard and with organization of the key comparison, F. G. Long and J. D. Kelley, for their construction of the two transfer standard packages, and S. Lu, for the development of data acquisition software and assistance with multiple calibrations of the transfer standards. The advice of C. R. Tilford throughout the course of this key comparison was also of great assistance. The many discussions with W. Guthrie of the Statistical Engineering Division at NIST concerning the statistical treatment of data were invaluable and are gratefully acknowledged as well.

REFERENCES

1. Comité International des Poids et Mesures (CIPM), *Mutual recognition of national measurement standards and of calibration and measurement certificates issued by national metrology institutes*, 1999, Paris, pp.45.

2. Calcatelli A., Molinar G. F., "The primary pressure scale in Italy from 10^{-6} Pa to 10^{9} Pa", in *Basic Metrology and Applications*, ed., Levrotto e Bella –Torino 1994.

3. Bergoglio M., Calcatelli A., "Vacuum measurement and traceability in Italy", *Proc. XIV IMEKO World Congress*, Tampere, June 1997.

4. Sharma J. K. N., Mohan P., "Use of successive expansion technique to determine volume ratios of ≈ 5000 for vacuum gauge calibration", *J. Vac. Sci. Technol.*, 1988, **A6** (6), 3148-3153.

5. Sharma J. K. N., Mohan P., Jitschin W., Rohl, P., "Intercomparison of vacuum standards between the Physikalisch Technische Bundensanstalt (Germany) and the National Physical Laboratory (India) using two spinning rotor gauges", *J. Vac. Sci. Technol*., 1989, **A7** (4), 2788-2793.

6. Sharma J. K. N., Mohan P., Sharma D. R., "Comparison of two primary pressure standards using spinning rotor gauges", *J. Vac. Sci. Technol.*, 1990, **A8** (2), 941-947.

7. Mohan P., Sharma D. R., Gupta A. C., "Comparison of an ultrasonic interferometer manometer and a static expansion system using a capacitance diaphragm gauge", *Metrologia*, 1996, **33**, 165-171.

8. Redgrave F. J., Forbes A. B., Harris P.M., "A discussion of methods for the estimation of volumetric ratios determined by multiple expansions", *Vacuum*, 1999, **53**, 159-162.

9. Elliott K. W. T., Clapham P. B., "The accurate measurement of the volume ratios of vacuum vessels", *NPL Report MOM 28*, January 1978.

10. Jitschin W., Migwi J.K., Grosse G., "Pressures in the high and medium vacuum range generated by a series expansion standard", *Vacuum*, 1990, **40**, 293-304.

11. Jitschin W., Migwi J.K., Grosse G., *Vacuum*, "Gauge calibration in the high and medium vacuum range by a series expansion standard", 1990, **41**, 1799.

12. Jousten K., Rupschus G., *Vacuum*, "The uncertainties of calibration pressures at PTB", 1993, **44**, 569-572.

13. Jousten K., Röhl P., Contreras V. A., *Vacuum*, "Volume ratio determination in static expansion systems by means of a spinning rotor gauge", 1999, **52**, 491-499.

14. Harrison E.R., Hatt D.J., Prowse D. B., Wilbur-Ham J., "A New Interferometric Manometer", *Metrologia*, 1976, **12**, 115-122.

15. Alasia F., Birello G., Capelli A., Cignolo G., Sardi M., "The HG5 laser interferometer mercury manometer of the IMGC", *Metrologia*, 1999, **36** (6), 499-503.

16. Hong S.S., Shin Y. H., Chung K. H., "Establishment of Low Vacuum Standards in the 1 Pa to 10^5 Pa Range", *J. Korean Vac. Soc*., 1996, **5** (3), 181-187.

17. Heydemann P. L. M., Tilford C. R., Hyland R. W., "Ultrasonic manometers for low and medium vacua under development at NBS", *J. Vac. Sci. Technol*., 1977, **14** (1), 597-605.

18. Tilford C. R., "Analytical procedure for determining lengths from fractional fringes", *Applied Optics*, 1977, **16** (7), 1857-1860.

19. Tilford C. R., Miiller A. P., Lu S., "A new low-range absolute pressure standard", *Proc. 1998 NCSL Workshop and Symposium*, National Conference of Standards Laboratories, Boulder, CO, USA, 245-256.

20. Tilford C. R., "The speed of sound in a mercury ultrasonic interferometer manometer", *Metrologia*, 1987, **24**, 121-131.

21. Miiller A. P., "Measurement performance of capacitance diaphragm gauges and alternative low-pressure transducers", *Proc. 1997 NCSL Workshop and Symposium*, National Conference of Standards Laboratories, Boulder, CO, USA, 287-299.

22. Miiller A. P., "Measurement performance of high-accuracy low-pressure transducers", *Metrologia*, 1999, **36** (6), 617-621.

23. Poulter K. F., Rodgers M.-J., Nash P. J., Thompson T. J., Perkin M. P., "Thermal transpiration correction in capacitance manometers", *Vacuum*, 1983, **33** (6), 311-316.

24. Takaishi T., Sensui Y., "Thermal transpiration effect of hydrogen, rare gases and methane", *Trans. Faraday Soc*., 1963, **59**, 2503-2514.

25. ISO, *Guide to the Expression of Uncertainty in Measurement* (International Organization for Standardization, Geneva, Switzerland, 1993).

26. Taylor B. N., Kuyatt C. E., *Guidelines for Evaluating and Expressing the Uncertainty of NIST Measurement Results*, NIST Technical Note 1297, 1994 Edition (U.S. Government Printing Office, Washington, 1994).

27. Youden W. J., "Graphical diagnosis of interlaboratory test results", *Industrial Quality Control*, 1959, **15** (11), 24-28.

28. Guthrie W. F., Statistical Engineering Division, National Institute of Standards and Technology, *Private Communication.*

29. Levenson M. S., Banks D. L., Eberhardt K. R., Gill L. M., Guthrie W. F., Liu H. K., Vangel M. G., Yen J. H., Zhang N. F., "An approach to combining results from multiple methods motivated by the ISO GUM", *J. Res. Natl. Inst. Stand. Technol*., 2000, **105** (4), 571-579.