CIPM Key Comparison

CCL-K2

Calibration of long gauge blocks

Final Report - Version B.5

Andrew Lewis, NPL, Teddington, March 2003

Contents

1 Introduction

The metrological equivalence of national measurement standards and of calibration certificates issued by national metrology institutes is established by a set of key comparisons chosen and organized by the Consultative Committees of the CIPM or by the regional metrology organizations in collaboration with the Consultative Committees.

At its meeting in September 1997, the Consultative Committee for Length, CCL, decided upon a key comparison on long gauge block measurements, numbered CCL-K2, starting in autumn 1999, with the National Physical Laboratory (NPL) as the pilot laboratory.

The results of this international comparison contribute to the Mutual Recognition Arrangement (MRA) between the national metrology institutes of the Metre Convention. This CIPM key comparison is linked with regional comparisons (RMO key comparisons) following exactly the same scheme. Laboratories participating in both the CIPM and the RMO comparisons establish the link between these comparisons and assure their equivalence.

2 Organisation

According to the rules set up by the BIPM [1] a small group from the list of participating laboratories drafted the detailed technical protocol. The group was composed of Jennifer Decker from the NRC, Canada, Nicholas Brown from CSIRO/NML, Australia and Andrew Lewis from the pilot laboratory, NPL. The protocol document and this report have been based on the corresponding documents for key comparison CCL-K1 [2]. The protocol document was issued to all participants at the start of the comparison.

2.1 Participants

All members of the CCL were invited to participate, subject to meeting certain technical requirements as laid out in the draft protocol document. In order to further reduce the number of participants to an acceptable level, each RMO was asked to limit the number of participants in their region, by its own decision process. This prevented the comparison taking too long with the commensurate risk of excessive damage to the artefacts. The participants were organized into regional groups to assist in the transportation of the artefacts (particularly as hand carriage of the items was specified in the protocol document). The list of participants is given in Table 1.

EUROMET Circulation		
Andrew Lewis	NPL (Pilot) Teddington Middlesex TW11 OLW United Kingdom	Tel.: +44 20 8943 6124 Fax : +44 20 8614 0533 e-mail: andrew.lewis@npl.co.uk
Alessandro Balsamo	IMGC-CNR Strada delle Cacce, 73 10135 Torino Italy	Tel.: +39 011 3977 470 Fax : +39 011 3977 459 e-mail: a.balsamo@imgc.to.cnr.it
Gerhard Bönsch	PTB Lab. 5.13 Postfach 3345 D-38023 Braunschweig Germany	Tel. +49 531 592 5130 Fax. +49 531 592 5015 e-mail: gerhard.boensch@ptb.de
SIM Circulation		
Hakima Belaidi	INMETRO Av. Nossa Senhora das Graças, 50 Xerém - 25250-020 Duques de Caxias Rio de Janeiro Brazil	Tel. +55 21 679 9271 or 9038 Fax +55 21 679 9207 or 1507 e-mail: hbelaidi@inmetro.gov.br
Jennifer Decker	Institute for National Measurement Standards National Research Council Canada Ottawa, K1A 0R6 Canada	Tel. +1 613 991 1633 Fax +1 613 952 1394 e-mail: jennifer.decker@nrc.ca

Theodore D. Doiron	National Institute of Standards and Technology Metrology (220), Room B118, NIST 100 Bureau Drive, Stop 8211 Gaithersburg, MD 20899-8211 USA	Tel. +1 301 975 3472 Fax +1 301 869 0822 e-mail: theodore.doiron@nist.gov
APMP Circulation		
Hirokazu Matsumoto	National Research Laboratory of Metrology NRLM 1-1-4 Umezono Tsukuba, Ibaraki 305-8563 Japan	Tel. +81 298 54 4034 Fax +81 298 54 4006 e-mail: hiro@nrlm.go.jp
Nicholas Brown	CSIRO/NML Division of Telecommunications and Industrial Physics PO Box 218 Lindfield, NSW 2070 Australia	Tel. +61 2 9413 7157 Fax +61 2 9413 7474 email: nick.brown@tip.csiro.au
Gao Sitian	National Institute of Metrology NIM No. 18, Bei San Huan Dong Lu Beijing 100013 China	Tel. +86 10 8425 1574 Fax +86 10 8425 1574 e-mail: gaost@nim.ac.cn
SADCMET Circulation		
Oelof Kruger	CSIR Division of Production Technology P.O. Box 395 Pretoria 0001 South Africa	Tel. +27 12 841 3005 Fax +27 12-841-2131 e-mail: oakruger@csir.co.za
COOMET Circulation		
Alexander Korolev	D.I.Mendeleyev Institute for Metrology VNIIM 19, perspective de Moskovsky 198005, St.Petersburg Russia	Tel. +7 812 251 8638 Fax +7 812 113 0114 e-mail: A.Korolev@vniim.ru
Roman Fira	Slovenský Metrologický Ústav SMU Karloveska 36 CS-842 55 Bratislava Slovakia	Tel. +421 7 602 94 284 Fax +421 7 654 29 592 e-mail: fira@smu.gov.sk

Table 1 Participating laboratories.

2.2 Schedule

The comparison has been carried out in a mixed form, circulation and star-type. After the standards had been circulated in a region, they went back to the pilot laboratory before being circulated in the next region. Because there was only one laboratory in the SADCMET region able to fulfil the technical requirements, they were included in the APMP circulation in order to provide a sufficient metrological link in their regional comparisons. The time schedule for the comparison is given in Table 2. Note that to take account of favourable travel opportunities, the measurements originally planned for September 2000 to November 2000 were delayed by one month. Note that due to scheduling requirements, it was necessary for the pilot laboratory to make its ‘official’ measurements part-way through the first circulation, and not at the very start of the comparison, as is usual. Also, re-scheduling of the COOMET regional circulation was necessary due to problems with temperature control during the first months of 2001.

Each laboratory was allowed one month in which to make its measurements and to prepare for transportation to the next participant. In cases where there was more than one month gap between participants, they were requested to only use one month for measurements. The schedule was designed to fit with the preferences of the laboratories for scheduling the measurements and any changes to the schedule, after the start of the circulation, were discussed and agreed among the participants and the CCL working group chairman.

RMO	Laboratory	Modified schedule	Date of measurement	Results received
EUROMET	IMGC	September 1999	September 1999	June 2000
	PTB	November 1999	November 1999	January 2000
	NPL	December 1999	January 2000	February 2000
Pilot Lab	NPL	January 2000	February 2000	N/A
SIM	NIST	March 2000	March 2000	June 2000
	INMETRO	April 2000	April 2000	June 2000
	NRC	May 2000	May 2000	August 2000
Pilot Lab	NPL	June 2000	June 2000	N/A
APMP	NRLM	July 2000	July 2000	September 2000
	NIM	August 2000	August 2000	February 2001
	CSIRO	October 2000	October 2000	January 2001
(SADCMET)	CSIR	November 2000	November 2000	December 2000
Pilot Lab	NPL	December 2000	December 2000	N/A
COOMET	SMU	March 2001	March – April 2001	May 2001
	VNIIM	May 2001	#June 2001	August 2001
Pilot Lab	NPL	May 2001	August 2001	N/A

Table 2 *Time schedule of the comparison. 'Modified schedule' refers to the latest schedule agreed among the participants. 'Results received' refers to the first date of receipt, by the pilot laboratory, of the official results of the participant (paper or electronic report).*

 # - delay in start of measurements at VNIIM due to problems clearing the gauges through customs, into Russia.

3 Standards

Four gauge blocks made of steel were circulated. The gauge blocks, which had been kindly donated by PTB, were selected as having a stable history of measurement and good flatness and variation in length. The gauge blocks were of rectangular cross section, according to international standard ISO 3650 (1998). The thermal expansion coefficient of the gauge blocks had been measured by the pilot laboratory and another laboratory (PTB) before the comparison. The weighted mean of the pilot laboratory and PTB results of expansion measurement (and their calculated uncertainties) were given to the participating laboratories in the technical protocol. The participating laboratories were only informed of the nominal length of the gauge blocks, as marked on their faces, the gauge material (steel), and the pre-determined expansion coefficients.

Serial number	**Nominal length (mm)**	α **(x 10^{-6} K^{-1})**	α **uncertainty (x 10^{-6} K^{-1})**
6071	175	10.968	0.016
6071	500	10.851	0.012
3701	500	10.624	0.012
3701	900	10.824	0.011

Table 3 *Standards used in the comparison. The uncertainties for the thermal expansion coefficients are given at k = 2.*

The standards were supplied in a custom made transport case, fashioned from aluminium and steel, containing high density foam, sculpted to make a tight fit with the gauge blocks, to prevent any motion of the gauge blocks and generation of excessive bending forces. The case was designed to be suitable for either cabin or hold transportation. The desire was for cabin transportation (hand carriage) with a fall-back option of transportation in the hold. The gauge blocks were accessible and visible with the lid opened and a pair of chamois gloves were included in case of any request by customs to handle the gauge bocks. The transport case and gauge blocks had a total mass in excess of 10 kg. Despite this being greater than the advertised cabin baggage allowance of most airlines, the airlines involved did not object to the hand carriage of the case in the aircraft cabin, provided they had been informed in advance.

Photo 1 *Gauge block transport case*

4 Measurement instructions and reporting of results

Before calibration, the gauge blocks had to be inspected for damage of the measurement surfaces. Any scratches, rusty spots or other damage had to be documented by a drawing using forms appended to the instructions.

The measurement quantity was the central length of the gauge blocks, as defined in International Standard ISO 3650. Any laboratory departing from the conditions specified in ISO 3650 had to make the relevant corrections to their measurand. ISO 3650 specifies that the gauge blocks had to be measured by interferometry, in the horizontal position wrung to a flat plate. The measurement result to be reported was the deviation of central length from nominal length, $\Delta l = l - L$. The results of the measurements on both sides (Δl_{left} and Δl_{right}) by wringing each measurement face in turn to the reference flat and the average of the two wringings had to be reported. The measurement results had to be appropriately corrected to the reference temperature of 20 °C using the thermal expansion coefficients given above. Additional corrections (aperture, phase correction) had to be applied according to the usual procedure of the laboratory. In cases where interferometry was not used, the participants were to interpret the instructions and reporting of results accordingly.

The uncertainty of measurement had to be estimated according to the *ISO Guide for the Expression of Uncertainty in Measurement*. In order to achieve optimum comparability, a mathematical model [3] containing the principal influence parameters for gauge block calibration by interferometry had been given in the technical protocols.

5 Measurement methods and instruments used by the participants

A wide variety of instruments and techniques were used to make measurements of the gauge blocks. The most important details of these instruments and techniques are reported in Tables 4 and 5. Of particular importance are the traceability route and the temperature of the gauge blocks during measurement. Participants were selected on the basis of their ability to measure at a temperature close to the reference temperature of 20 °C so as to minimize the uncertainty due to the thermal expansion correction. Also, independent traceability to the realisation of the metre was required as a condition of participation.

The majority of the participants used some type of direct interferometry for the measurement of at least the shorter gauge blocks (up to 500 mm). These interferometers show a direct view of the gauge block and platen surface with a superimposed fringe pattern. Various techniques were used to measure these fringe patterns to determine the fringe shift on the gauge surface relative to the surface of the platen.

One participant used an instrument designed for the measurement of linescales, to measure the 900 mm gauge block. In order to make a measurement, small fused silica blocks were wrung to the ends of the gauge block, and the distance between chrome lines on the surfaces of the blocks was measured. Then the blocks were wrung together and the separation of the lines measured again, in order to provide a correction for one wringing film thickness.

In all instruments, determination of the refractive index is important. Two techniques were used: direct evaluation of the refractive index by use of an *in situ* refractometer; and calculation of the refractive index based on measurements of air parameters such as temperature, pressure, humidity, and carbon dioxide content, and use of empirical equations.

Lab.	Type of equipment	Traceability route	Measurement method	Platen material	Temperature / °C
IMGC	SIP CLP10A interferometric comparator	Frequency stabilized laser (633 nm), calibrated by iodine stabilized laser.	Travelling carriage on comparator, measured interferometrically. Fringe fraction evaluation by eye on magnified screen, with white light interferometry used as fiducial indicator.	Steel	19.911 to 20.060
PTB	Kösters-Zeiss interferometric comparator.	Directly via use of iodine stabilized lasers at 633 nm, 612 nm, 515 nm.	Direct interferometry using exact-fractions. [4]	Steel	19.997 to 20.038
NPL	NPL design Twyman-Green phase stepping interferometer.	Frequency stabilized lasers, 633 nm, 543 nm, 612 nm calibrated by iodine stabilized lasers. 633 nm wavelength is reference.	Phase-stepping fringe fractions measurement using direct interferometry at three wavelengths. Method of excess-fractions, basing the result on the red wavelength. Refractive index determination via air temperature, pressure, humidity and CO_2 measurements. [5]	Steel	20.033 to 20.084
NIST	Zeiss gauge block interferometer. Linescale measuring instrument	Frequency stabilized laser (633 nm), calibrated by iodine stabilized laser.	Interference pattern photographed then analysed using a calibrated travelling microscope, to measure the fringe fractions. 900 mm gauge measured in linescale instrument [6], with fused silica end blocks wrung to the gauge – separation of chrome lines on blocks.	Steel. (Fused silica blocks, used in linescale equipment).	19.900 to 20.100 (175-500 mm) 19.995 to 20.005 (900 mm)
INMETRO	Kösters-Zeiss interferometric comparator.	Directly via Kr radiations, 642 nm, 605 nm, 565 nm. 605 nm wavelength is reference.	Direct interferometry using exact-fractions. Comparative measurement for the 900 mm gauge block, based on a shorter block.	Steel	19.90 to 19.99
NRC	NRC design Twyman-Green interferometer.	Frequency stabilized lasers, 633 nm, 543 nm, 612 nm calibrated by iodine stabilized lasers. 633 nm wavelength is reference.	Method of excess-fractions. Refractive index determination via air temperature, pressure and humidity measurements. [7, 8]	Steel	19.989 to 20.006
NRLM	NRLM design Twyman-Green interferometer.	Directly via use of iodine stabilized lasers at 633 nm and 532 nm.	Computer assisted interpretation of video image.	Steel	19.7 to 20.2
NIM	Modified Zeiss (175 mm) and Kösters interferometers.	Lamb-dip stabilized He-Ne laser at 633 nm (175 mm) or Kr lamp.	Direct interferometry –fringe fraction determination by eye or automatically (175 mm).	Glass	19.9 to 20.1
CSIRO	CSIRO modified Michelson and Kösters interferometers.	Directly via use of iodine stabilized lasers at 633 nm, 612 nm, 543 nm.	Method of excess-fractions. Refractive index determination via air temperature, pressure and humidity measurements.	Steel	19.882 to 20.098
CSIR	CSIR design moving probe system with interferometer distance measurement.	Frequency stabilized laser (633 nm), calibrated by iodine stabilized laser.	Probe with plane mirrors attached, contacts both faces of gauge block. Probe diameter determined using measurement of short gauge blocks.	N/A	19.981 to 20.098
SMU	SMU design Michelson interferometer.	Frequency stabilized laser (633 nm), calibrated by iodine stabilized laser.	Dynamic fringe counting and interpolation, using white light as fiducial indicator. Refractive index determination via air temperature, pressure, humidity measurements.	Steel	20.022 to 20.104 (corrected)
VNIIM	VNIIM design gauge block interferometer.	633 nm He-Ne laser.	Direct interferometry using fringe fractions and white light interferometry for fringe order determination.	Steel	19.9 to 20.1

Table 4 Measurement instruments and conditions reported by the participating laboratories.

Lab.	Platen weight compensation method	Phase/roughness correction determination method	Phase correction value(s)	Vertical to horizontal correction	Measurement position (Vertical, Horizontal) and temperature during measurement (°C)				Refractive index determination method
					175 mm S/N 6071	500 mm S/N 6071	500 mm S/N 3701	900 mm S/N 3701	
IMGC	Counterbalance	Not measured	-	-	H 19.977 to 20.087	H 19.933 to 20.055	H 19.971 to 20.060	H 19.911 to 19.977	Birch & Downs
PTB	Move supports [12]	TIS	-1 to –4 nm	-	H 20.001 to 20.007	H 19.998 to 20.015	H 20.002 to 20.015	H 20.020 to 20.038	Vacuum cell
NPL	Move supports [15]	Not measured	-	-	H 20.040 to 20.048	H 20.031 to 20.042	H 20.036 to 20.081	H 20.034 to 20.038	Birch & Downs
NIST	-	Not specified	Not specified	Not specified	V 19.9 to 20.1	V 19.9 to 20.1	V 19.9 to 20.1	H 19.995 to 20.005	Edlen
INMETRO	Move supports	Not measured	-	-	H 19.912 to 19.991	H 19.919 to 19.988	H 19.934 to 19.964	H 19.899 to 19.956	Vacuum cell
NRC	Move supports [15]	Not measured	-	-	H 19.991 to 19.999	H 19.993 to 19.997	H 19.992 to 20.006	H 19.989 to 19.999	Ciddor
NRLM	Counterbalance	Stack	-6 nm	-	H 19.890	H 19.900	H 19.900	H 19.930	Ciddor
NIM	Move supports	Stack (175 mm)	+60 nm	60 nm (175 mm)	V 19.940	H 20.000	H 19.990	H 19.970	Not specified
CSIRO	Counterbalance	Not measured	-	$0.186 \times 10^{-6} L^2$ 6 nm (175 mm) 47 nm (500 mm)	V 19.936 to 20.063 H: 19.836	H 19.835 to 20.013 V: 20.035	H 20.060 to 20.071 V: 20.139	H 19.927 to 20.007	Ciddor
CSIR	-	-	-	-	H 19.981 to 20.098	H 19.981 to 20.098	H 19.981 to 20.098	H 19.981 to 20.098	Edlen
SMU	Move supports	Not measured	-	-	H 20.022 to 20.104	H 20.022 to 20.104	H 20.022 to 20.104	H 20.022 to 20.104	Edlen
VNIIM	Move supports	Not specified	Not specified	-	H 19.888 to 20.080	H 19.758 to 19.884	H 19.956 to 19.996	H 19.872 to 20.082	Not specified

Table 5 Additional measurement conditions and details reported by the participating laboratories.

<u>Refractive index determination method:</u>
Edlen = The refractive index of air [9]
Birch & Downs = Correction to the updated Edlén equation for the refractive index of air [10]
Ciddor = Refractive index of air: new equations for the visible and near infrared [11]
Vacuum cell = Internal refractometer or vacuum cell for absolute refractive index determination

<u>Phase/roughness method</u>
TIS = Total Integrated Scatter method using integrating sphere [13, 14]
Stack = Traditional ‘stack’ or ‘pack’ method based on gauges wrung separately and as a stack

6 Stability and condition of the gauge blocks

NPL and PTB made interferometric calibrations before the start of the comparison (Prelim) and PTB (who donated the gauge blocks) had some prior information concerning the historic stability of the standards. The pilot laboratory made interferometric calibrations at the end of the EUROMET calibration (its own official measurements), between the SIM and APMP circulations (Interim1), between the APMP and COOMET circulations (Interim2), and at the end of the circulation (Final) using the same equipment as used to perform its 'official' calibrations. These interim calibrations included measurement of the central length, flatness and variation in length. The same platens were used for all of these measurements by the pilot laboratory, with each gauge block being wrung to the same platen for all of its measurements.

6.1 Central length stability

Figures 1(a) through 1(d) show the measurements of the pilot laboratory used to verify the stability of the gauge blocks' central length.

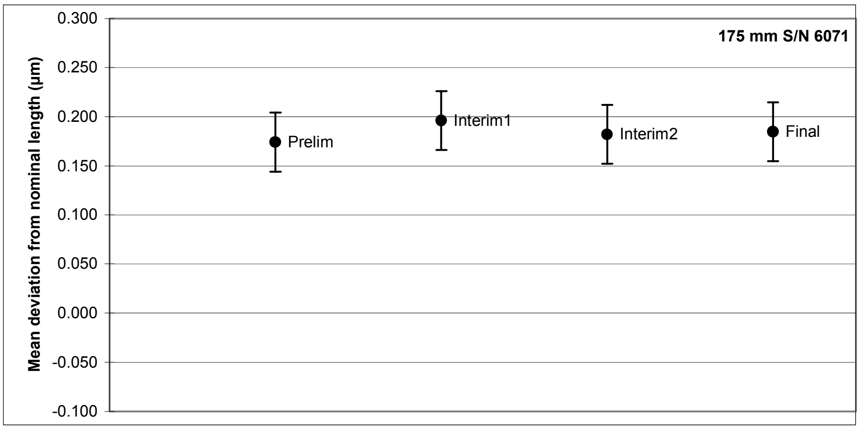


Figure 1(a) Stability of 175 mm gauge block (S/N 6071) during comparison: interim length measurements of the pilot laboratory. Uncertainty bars show standard uncertainty (k=1).

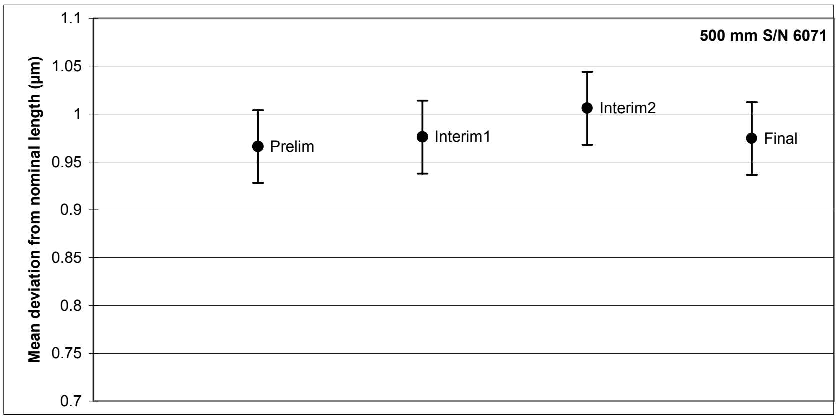


Figure 1(b) Stability of 500 mm gauge block (S/N 6071) during comparison: interim length measurements of the pilot laboratory. Uncertainty bars show standard uncertainty (k=1).

The uncertainty bars in Figures 1(a) through 1(d) are standard uncertainties of the pilot laboratory's usual measurement technique. Because the same equipment, platens, operator and technique, were used for these measurements, several uncertainty sources will be correlated for the four measurements (*e.g.* phase correction uncertainty) and so in terms of possible changes in length, the uncertainties would be somewhat reduced. The measured overall changes in length for the gauge blocks in Figures 1(a) through 1(d) were: 22 nm, 40 nm, 27 nm and 38 nm, respectively.

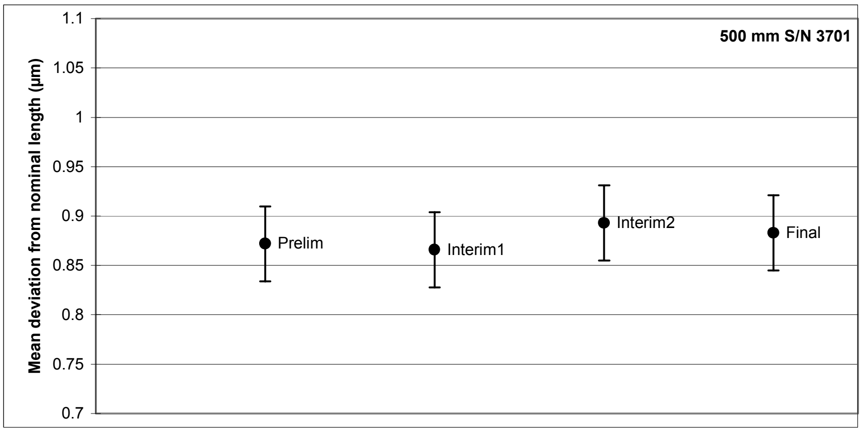


Figure 1(c) *Stability of 500 mm gauge block (S/N 3701) during comparison: length measurements of the pilot laboratory. Uncertainty bars show standard uncertainty (k=1).*

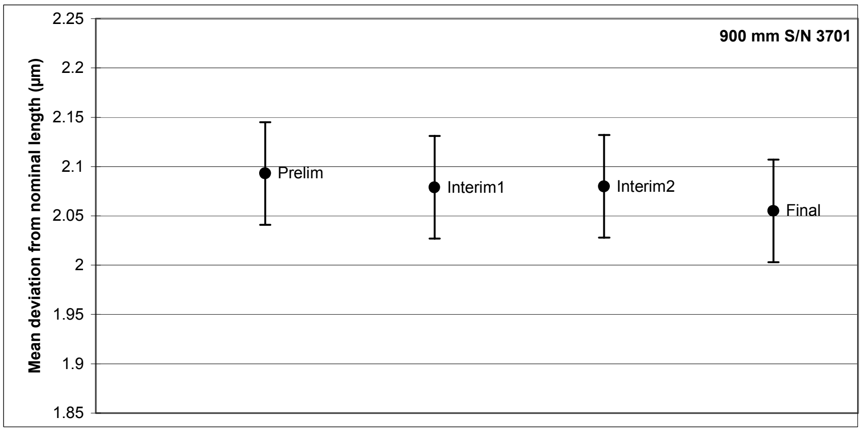


Figure 1(d) *Stability of 900 mm gauge block (S/N 3701) during comparison: length measurements of the pilot laboratory. Uncertainty bars show standard uncertainty (k=1).*

6.2 Condition of the gauge blocks

The gauge blocks were essentially free of any damage at the beginning of the comparison. The participating laboratories were asked to document any scratches and other damages on the measurement surfaces by a drawing to be made when receiving the gauge blocks. As the comparison progressed, more scratches appeared on the measurement surfaces of the gauge blocks as well as some marks on the side faces. Some indentations became apparent on the narrow faces close to the Airy points. Copies of the drawings of the measurement faces that were supplied by the participants may be found below (Figures 2(a) and 2(b)). Note that the scanning of the pictures and their transmission to the pilot laboratory has, for some participants, introduced a change in contrast to their pictures.

It is interesting to note the different interpretations of the gauge block surface condition, as reported by the participants in their drawings of the gauge block surfaces. There are clearly some surface defects which are reported by several participants, whereas other defects which are reported by one participant were not reported by later participants. The pilot laboratory is in a unique position of seeing the gauge blocks several times throughout the comparison, as well as seeing the individual reports of the participants. This gives the pilot laboratory the ability to more accurately monitor the damage to the gauge blocks. However the prevalence of deeper scratches later in the comparison may mask the lighter scratching which was apparent at the start, leading to differences of opinion on the surface quality. IMGC did not submit a report, as they found no scratches worthy of note, when they examined the gauge blocks at the start of the comparison.

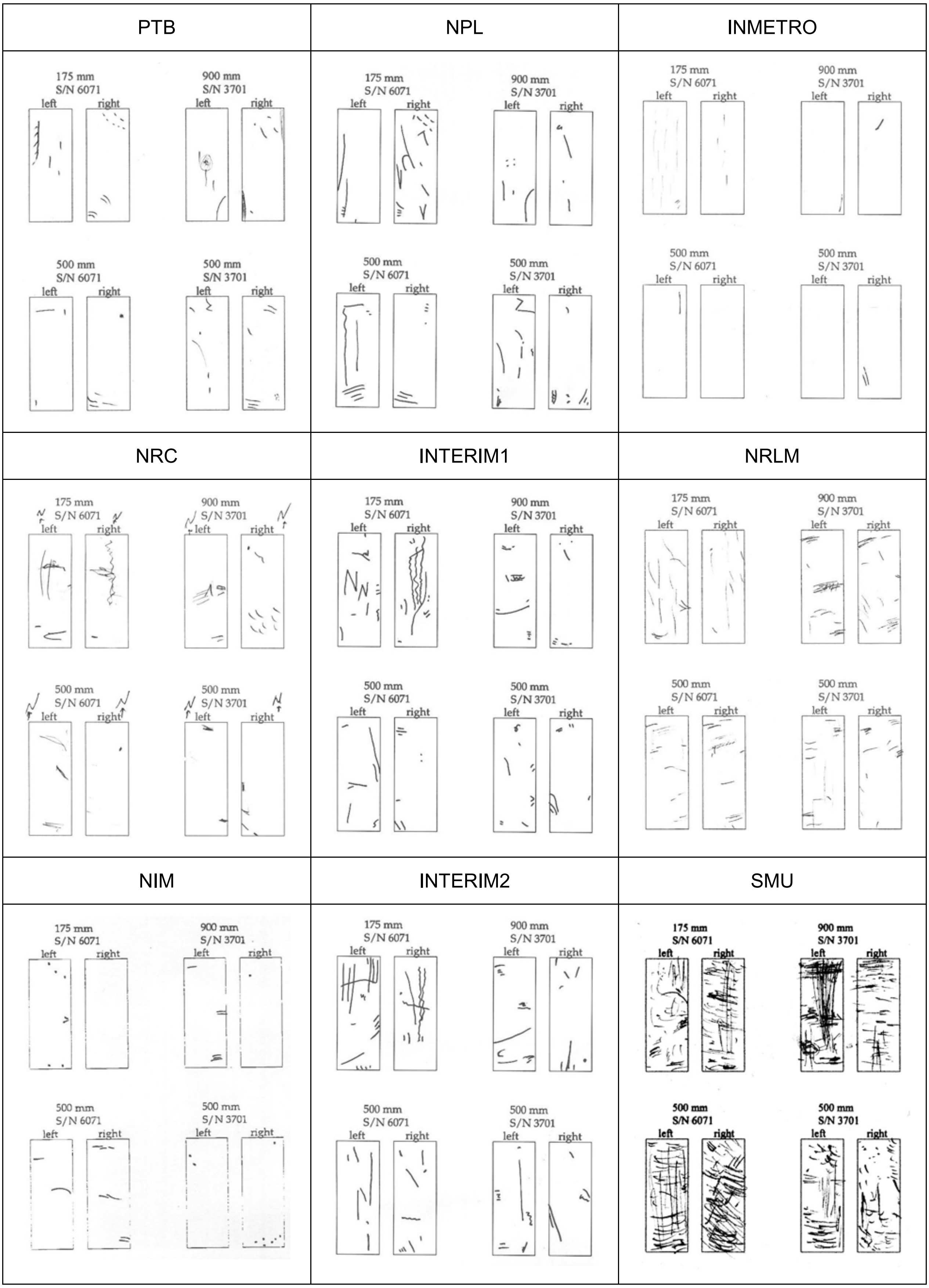
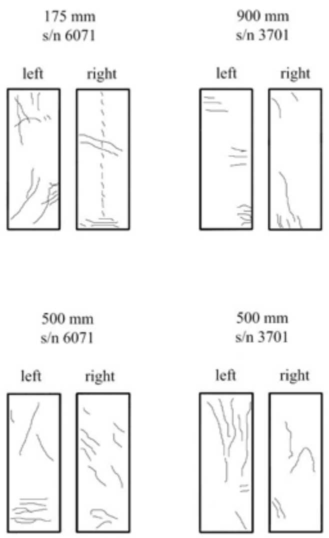


Figure 2(a) Gauge condition reports received from participants

VNIIM



FINAL (PILOT)

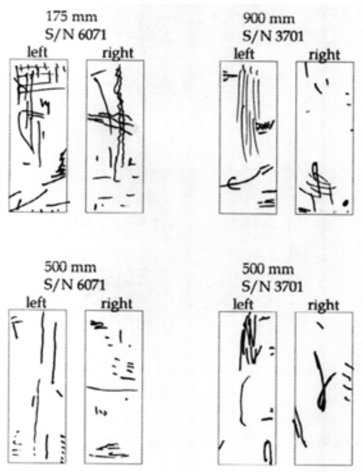


Figure 2(b) Gauge condition reports received from participants

6.3 Stability of flatness and variation in length

Although the participants were not required to measure the flatness or variation in length of the gauge blocks, the pilot laboratory made measurements of these parameters at the start, middle and end of the comparison to check for stability of the gauge blocks. This is important as any change to these values may have an effect on the central length measurement, depending on the measurement technique used.

The pilot laboratory measurements of flatness and variation in length were performed at the same time as the re-measurements of the central length. The measurements were made using a phase stepping technique with an estimated (2 sigma) uncertainty of 30 nm.

Table 6 shows the stability data for the flatness and variation in length, as measured by the pilot laboratory.

	175 mm, S/N 6071				**500 mm, S/N 6071**			
	Flatness (nm)		Variation (nm)		Flatness (nm)		Variation (nm)	
	Right	Left	Right	Left	Right	Left	Right	Left
Start	49	-	130	-	54	56	180	166
Interim1	62	54	121	75	62	80	210	252
Interim2	51	55	76	59	44	60	239	239
Final	45	60	105	74	55	60	218	260
	500 mm, S/N 3701				**900 mm, S/N 3701**			
	Flatness (nm)		Variation (nm)		Flatness (nm)		Variation (nm)	
	Right	Left	Right	Left	Right	Left	Right	Left
Start	44	38	132	128	123	112	340	179
Interim1	38	49	110	121	134	142	270	210
Interim2	35	42	144	108	104	110	256	218
Final	47	60	157	117	123	173	323	215

Table 6 Stability of gauge flatness and variation in length, as determined by the pilot laboratory. Right and left refer to the face which is wrung (corresponding to the length determination, not to the face which is visible). Uncertainty is estimated to be ~30 nm (k=2) for each result.

Generally, the flatness and variation data remained almost unchanged throughout the comparison, indicating that the gauge blocks were quite stable. The most stable of the gauge blocks were the 500 mm gauge block, S/N 3701, and the 175 mm gauge block, where there was almost no change in flatness and less than 50 nm change in variation in length. The other 500 mm gauge block (S/N 6071) showed a 94 nm change in the variation in length and the 900 mm gauge block exhibited an 84 nm change in variation in length and a 63 nm change in flatness.

A change in the variation in length or flatness of the end faces may have a direct influence on the measured central length, and so these changes should be accounted for in the uncertainty estimation. Of course, no single participant can observe such changes, since they only make one measurement, and so the uncertainty due to change in artefact geometry is an uncertainty of the artefact, rather than of the participants' measurement processes. This will be considered in a later section.

7 Measurement results, as reported by participants

7.1 Deviation from nominal length

In Tables 7(a) through 7(d), all measurement results for the deviation from nominal length for the four gauge blocks are given along with their combined standard uncertainties, as reported by the participants. Results reported are the central deviation from nominal length with the left face wrung (ΔL_{left}),central deviation from nominal length with the right face wrung (ΔL_{right}), and the mean of these results (ΔL_{mean}), for each gauge block, for each laboratory. The standard (k = 1) uncertainty reported by each laboratory is also given. Note that NIST was not able to measure the gauge blocks with the right face wrung due to poor wringing quality, so NIST only reported a value for ΔL_{left}, which is therefore also taken as the value for ΔL_{right} and ΔL_{mean}. For laboratories which did not make direct interferometric measurements, the terms *left* and *right* refer to the gauge blocks turned end for end between two sets of measurements. All reported data is rounded to the nearest nanometre.

175 mm S/N 6071				
LAB	**ΔL_{left} (µm)**	**ΔL_{right} (µm)**	**ΔL_{mean} (µm)**	**u_c (µm)**
IMGC	0.150	0.130	0.140	0.028
PTB	0.121	0.123	0.122	0.013
NPL	0.162	0.160	0.161	0.030
NIST	0.142	0.142	0.142	0.016
INMETRO	0.150	0.150	0.150	0.020
NRC	0.123	0.127	0.125	0.027
NRLM	0.144	0.153	0.148	0.019
NIM	0.195	0.194	0.194	0.019
CSIRO	0.157	0.149	0.154	0.023
CSIR	0.200	0.160	0.180	0.110
SMU	0.410	0.420	0.410	0.038
VNIIM	0.307	0.316	0.312	0.021
Mean	0.188	0.185	0.187	0.030

Table 7(a) Results for the 175 mm gauge block, S/N 6071.

500 mm S/N 6071				
LAB	**ΔL_{left} (µm)**	**ΔL_{right} (µm)**	**ΔL_{mean} (µm)**	**u_c (µm)**
IMGC	0.915	0.917	0.916	0.033
PTB	0.921	0.908	0.915	0.016
NPL	0.963	0.961	0.962	0.038
NIST	0.908	0.908	0.908	0.023
INMETRO	0.920	0.940	0.930	0.020
NRC	0.885	0.877	0.881	0.067
NRLM	0.934	0.943	0.938	0.039
NIM	1.010	1.004	1.007	0.060
CSIRO	0.870	0.915	0.885	0.050
CSIR	0.950	1.010	0.980	0.150
SMU	1.330	1.310	1.320	0.074
VNIIM	0.935	0.949	0.952	0.056
Mean	0.962	0.970	0.966	0.052

Table 7(b) Results for the 500 mm gauge block, S/N 6071.

500 mm S/N 3701				
LAB	**ΔL $_{left}$ (μm)**	**ΔL $_{right}$ (μm)**	**ΔL $_{mean}$ (μm)**	**u_c (μm)**
IMGC	0.806	0.822	0.814	0.033
PTB	0.807	0.807	0.807	0.015
NPL	0.852	0.870	0.861	0.038
NIST	0.781	0.781	0.781	0.023
INMETRO	0.830	0.830	0.830	0.020
NRC	0.785	0.788	0.786	0.066
NRLM	0.867	0.850	0.858	0.039
NIM	0.910	0.914	0.912	0.060
CSIRO	0.808	0.835	0.818	0.050
CSIR	0.880	0.850	0.870	0.150
SMU	1.330	1.190	1.250	0.072
VNIIM	0.866	0.870	0.868	0.056
Mean	0.877	0.867	0.871	0.052

Table 7(c) Results for the 500 mm gauge block, S/N 3701.

900 mm S/N 3701				
LAB	**ΔL $_{left}$ (μm)**	**ΔL $_{right}$ (μm)**	**ΔL $_{mean}$ (μm)**	**u_c (μm)**
IMGC	2.028	2.038	2.033	0.042
PTB	1.985	1.981	1.983	0.021
NPL	2.081	2.032	2.057	0.052
NIST	2.075	2.075	2.075	0.060
INMETRO	2.040	2.000	2.020	0.035
NRC	2.026	1.981	2.004	0.118
NRLM	2.087	2.052	2.070	0.068
NIM	2.173	2.146	2.160	0.136
CSIRO	2.000	1.963	1.982	0.087
CSIR	2.070	1.950	2.010	0.250
SMU	2.910	2.890	2.900	0.125
VNIIM	2.160	2.165	2.165	0.100
Mean	2.136	2.106	2.122	0.091

Table 7(d) Results for the 900 mm gauge block, S/N 3701.

Figures 2(a) through 2(d) show the ΔL_{mean} results with standard uncertainty bars.

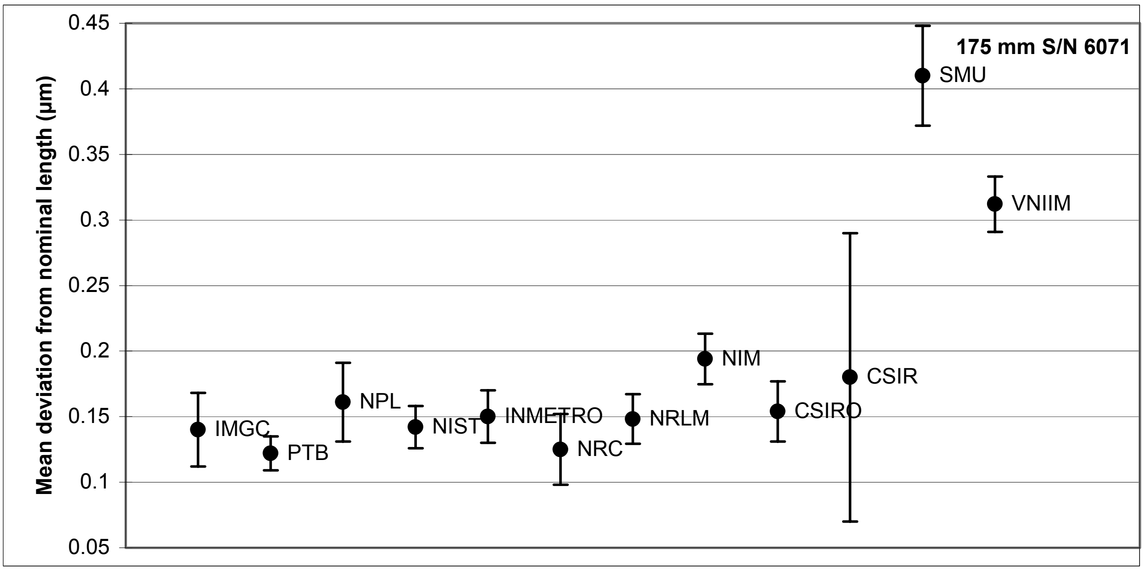


Figure 3(a) Results for the 175 mm gauge block, S/N 6071 (standard uncertainty bars shown).

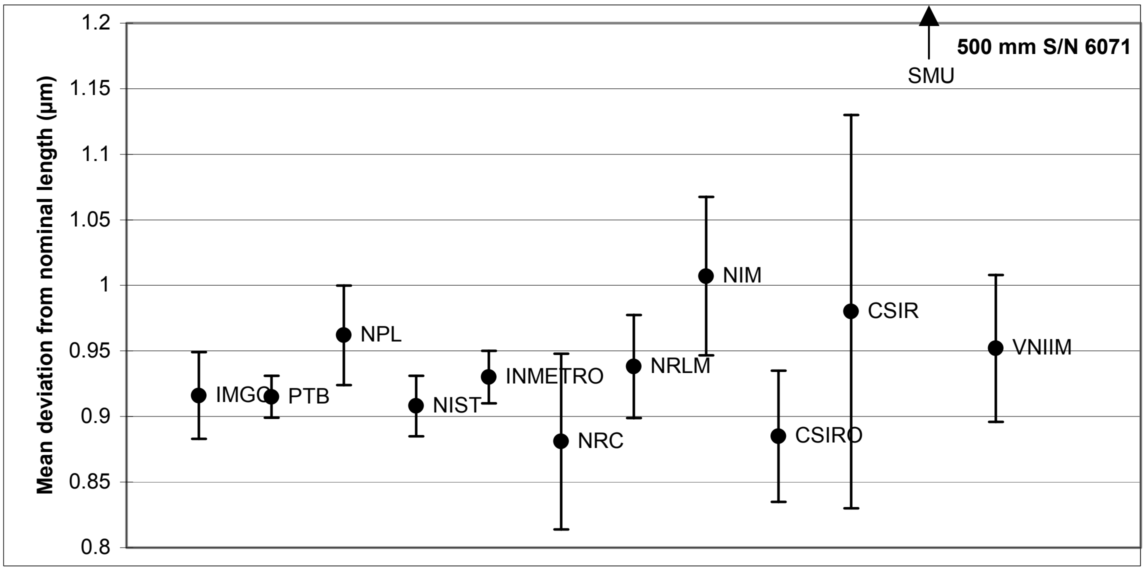


Figure 3(b) *Results for the 500 mm gauge block, S/N 6071 (standard uncertainty bars shown).*

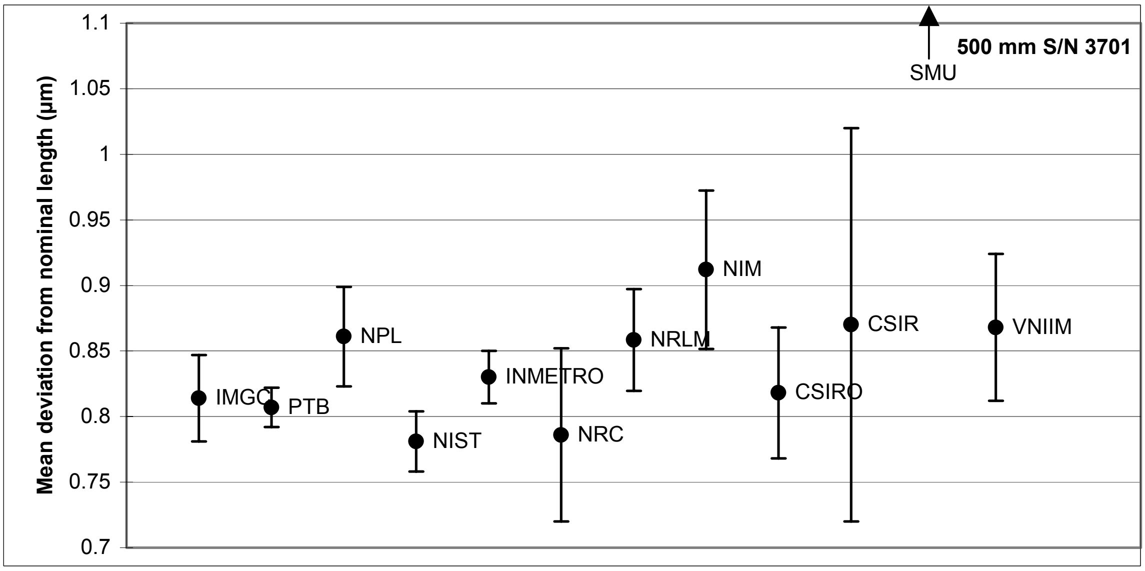


Figure 3(c) *Results for the 500 mm gauge block, S/N 3701 (standard uncertainty bars shown).*

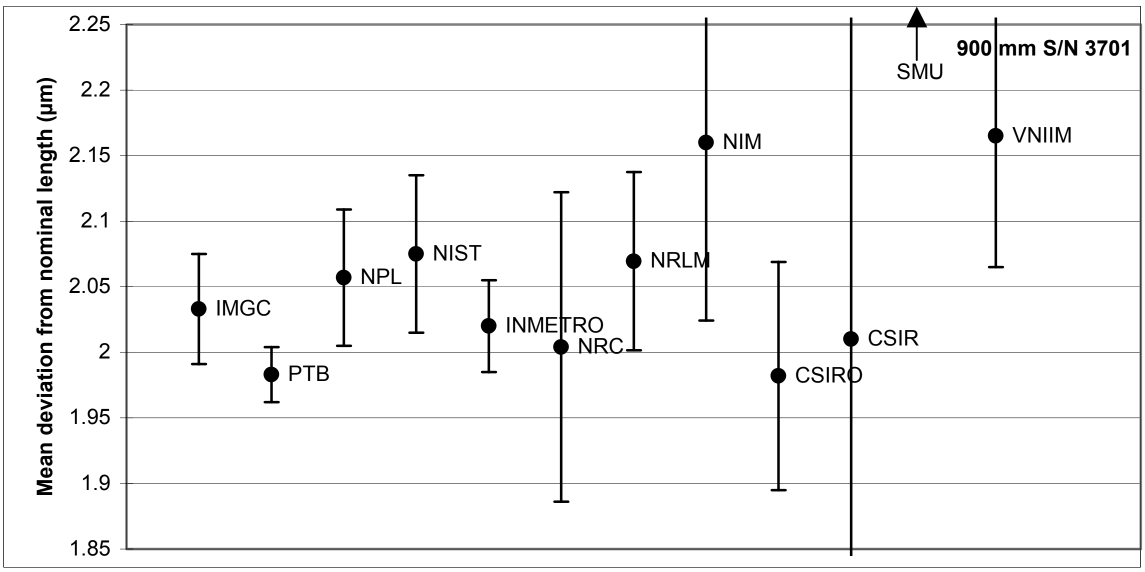


Figure 3(d) *Results for the 900 mm gauge block, S/N 3701 (standard uncertainty bars shown).*

7.2 Difference between left and right wringing

The laboratories were requested to measure the gauge blocks with both the left and the right measurement surface wrung and to report the results from the individual wringings and the mean. Table 8 shows the differences between the two wringings of all laboratories for all four gauge blocks, separately. Note that one laboratory was not able to perform measurements with both faces independently wrung, (NIST) so this laboratory does not report results in Table 8, or Figure 4.

LAB	175 mm, SN 6071	500 mm, SN 6071	500 mm, SN 3701	900 mm, SN 3701
IMGC	20	-2	-16	-10
PTB	-2	13	0	4
NPL	2	2	-18	49
INMETRO	0	-20	0	40
NRC	-4	8	-3	45
NRLM	-9	-9	17	36
NIM	1	6	-4	27
CSIRO	8	-45	-27	37
CSIR	40	-60	30	120
SMU	-10	20	140	20
VNIIM	-9	-14	-4	-5
Mean	3	-9	10	33

Table 8 Differences between left and right wringing for all four gauge blocks ($\Delta L_{left} - \Delta L_{right}$), with the result given in nm.

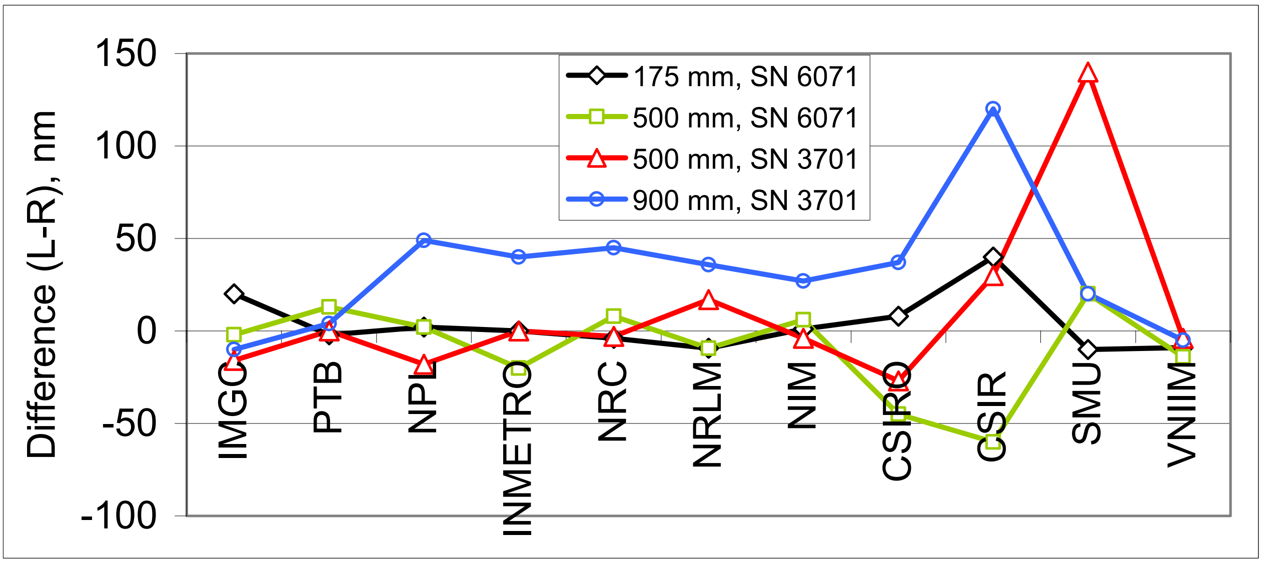


Figure 4 Differences between length measurements for left and right wringing ($\Delta L_{left} - \Delta L_{right}$).

Table 8 and Figure 4 indicate that the difference in length between left and right wringing increases with gauge block length, with the 900 mm gauge block exhibiting the largest mean difference.

8 Measurement uncertainties

8.1 Model equations

The participants were asked in the technical protocol of the comparison to estimate the uncertainty of measurement according to the *ISO Guide for the Expression of Uncertainty in Measurement*. An example mathematical model [3] was given and the participants were encouraged to use this model as closely as possible to allow for a detailed comparison of all the uncertainties. However, since a variety of measurement techniques have been used, in several cases the uncertainties are not given according to this model.

The uncertainties in the determination of the following model parameters were taken into account by the majority of the participants:

λ_i vacuum wavelengths of the different light sources used;

F_i fractional part of fringe order;

n index of refraction of the air;

Δt_G difference of the gauge block temperature from the reference temperature of 20 °C;

α linear coefficient of thermal expansion of the gauge block;

δl_Ω obliquity correction for the shift in phase resulting from the angular alignment errors of the collimating assembly;

Δl_s aperture correction accounting for the shift in phase resulting from the finite aperture diameter *s* of the light source;

δl_A correction for wave front errors as a result of imperfect interferometer optics;

δl_G correction accounting for flatness deviation and variation in length of the gauge block;

δl_w length attributed to the wringing film;

Δl_Φ phase change accounting for the difference in the apparent optical length to the mechanical length;

In Table 9, the uncertainty contributions are summarized for all laboratories. The numerical values are standard uncertainties given in nanometres. The length dependent terms are written in italic letters and were calculated for a gauge block length of 500 mm. In the last row, the combined standard uncertainty has been calculated by a simple quadratic sum. This might not necessarily be identical to the combined uncertainty quoted by the laboratory for the 500 mm gauge blocks, because they might have used further contributions, correlations and second order terms, which are not given in the table.

For some laboratories (NIST, CSIR) the quoted uncertainties did not align exactly with the model equation. In these cases, all uncertainties have been introduced into Table 9, by attributing them to the nearest meaningful uncertainty parameter.

Model parameter	IMGC	PTB	NPL	NIST	INMETRO	NRC	NRLM	NIM	CSIRO	CSIR	SMU	VNIIM
λ_i	*2*	*0.08*	*5*	*1.5*	*1.6*	*2.5*	*0.01*	*12.8*	*0.5*	*5*	*7*	*2.5*
F_i	18	5	0.9	3	5	2	3.2	12.5	2	-	12	5.6
n	*14*	*5*	*18.2*	*15.1*	*6*	*12*	*17*	*24*	*22.5*	*50*	*15*	*25*
Δt_G	*12.5*	*5.5*	*14.5*	*6*	*7.4*	*44*	*32.6*	*44.5*	*42*	*57.5*	*66*	*28*
α	*0.5*	*0.1*	*0.4*	-	*0.3*	*0.2*	*0.5*	*2.3*	*0.05*	*2.3*	*0.13*	*25*
δl_Ω	*0*	*1.5*	*0.8*	*0*	*0.6*	*50*	*2.2*	*16.9*	*5*	*120*	*4*	*23*
Δl_s	-	*0.5*	*0*	*0*	*0.5*	*0.5*	*0.1*	*11.2*	*1.5*	-	*0.2*	*23*
δl_A	4	5	0.5	0	3	3	6.7	6.7	4	-	5	3
δl_G	1	6	20	8	11	6	3	13.4	6	20	5	3
δl_W	10	9	5	10	6	10	10	7.6	6	-	6	3
Δl_Φ	17	5.8	20	7.2	6	10	5	7.1	11	-	15	3
Δl_h	-	-	-	-	-	-	-	4.9	-	-	-	-
u_c	**34**	**17**	**38**	**23**	**18**	**70**	**40**	**61**	**50**	**144**	**72**	**56**

Table 9 *Standard uncertainties (in nm) quoted for a 500 mm gauge block, by the different laboratories for the uncertainty contributions given in the model of the technical protocols, and combined uncertainty calculated from these values. The quadrature summed standard uncertainties (u_c) have been rounded up to the nearest nm.*

Δl_h represents the 'correction for the state where the gauge block is lying in a vertical plane', as reported by one participant.

Examination of Table 9 shows that the uncertainty due to the coefficient of thermal expansion, α, is very small, compared to other sources of uncertainty. This is presumably due to *a priori* knowledge of these coefficients, through their determination by the pilot laboratory and another laboratory before the start of the comparison. In hindsight, this allows laboratories with poor temperature control to achieve uncertainties which are better than would normally be achievable with 'unknown' customer gauges.

The largest source of uncertainty is due to the uncertainty in the determination of the gauge block temperature. This uncertainty is multiplied by the gauge block coefficient of thermal expansion which is approximately 10^{-5} K^{-1}. In order, therefore, to reduce this uncertainty below 10 nm (for a 500 mm gauge block) requires temperature measurement with an uncertainty below 2 mK, which is quite difficult to achieve.

9 Analysis of the reported results

The reported measurement results are now analysed by simple statistical means to allow identification of any significant bias or outliers, and to investigate the statistical distribution of the results. A quantitative analysis of the deviations of the results of each laboratory from some reference value can only be made once the key comparison reference values have been determined and confirmed by the CCL.

From Tables 7(a) though 7(d) and Figures 3(a) through 3(d) it is clear that the uncertainties quoted by the participants are different from one participant to another, and that the uncertainties depend on the length of the gauge block being measured. Thus analysis via use of the simple arithmetic mean as an estimator of the true mean is not suitable and instead, the weighted mean should be used. This approach requires that the participants have made correct estimates of their uncertainty of measurement otherwise a too low uncertainty will place undue emphasis on the result of that particular laboratory.

9.1 Derivations

For each laboratory, i, which measures each gauge block, j, let the measured deviation from nominal size (after making all required corrections) be denoted x_{ij}. The number of laboratories, I, is 12 and the number of gauge blocks, J, is 4. Since the four gauge blocks are four physically different length artefacts with four different lengths, thermal expansion coefficients, material properties *etc*, it is reasonable to expect that the data x_{ij} come from four separate populations (one per gauge block) and so analysis should be on a gauge-by-gauge basis.

Thus, for a particular gauge block, j :

Each laboratory reports a measured value, x_i , and its associated standard uncertainty $u(x_i)$.

The normalised weight, w_i , for the result x_i is given by:

$$w_i = C \cdot \frac{1}{[u(x_i)]^2} \tag{1}$$

where the normalising factor, C, is given by:

$$C = \frac{1}{\sum_{i=1}^{I} \left(\frac{1}{u(x_i)} \right)^2} \tag{2}$$

Then the weighted mean, $\bar{x}_w$, is given by:

$$\bar{x}_w = \sum_{i=1}^{I} w_i \cdot x_i \tag{3}$$

The uncertainty of the weighted mean can be calculated as either the so-called internal or external standard deviation, $u_{int}(\bar{x}_w)$ and $u_{ext}(\bar{x}_w)$, respectively. The internal standard deviation is based on the estimated uncertainties $u(x_i)$ as reported by the participants, whereas the external standard deviation is the standard deviation of the spread of the actual results, x_i, weighted by the uncertainties $u(x_i)$:

$$u_{int}(\bar{x}_w) = \sqrt{\frac{1}{\sum_{i=1}^{I} \left(\frac{1}{u(x_i)} \right)^2}} = \sqrt{C} \tag{4}$$

$$u_{ext}(\bar{x}_w) = \sqrt{\frac{1}{(I-1)} \cdot \frac{\sum_{i=1}^{I} w_i (x_i - \bar{x}_w)^2}{\sum_{i=1}^{I} w_i}} \qquad (5)$$

Substituting (1) into (5) gives:

$$u_{ext}(\bar{x}_w) = \sqrt{\frac{1}{(I-1)} \cdot \frac{\sum_{i=1}^{I} \frac{1}{[u(x_i)]^2} (x_i - \bar{x}_w)^2}{\sum_{i=1}^{I} \frac{1}{[u(x_i)]^2}}} \qquad (6)$$

After deriving the weighted mean and its associated uncertainty, the deviation of each laboratory's result from the weighted mean is determined simply as $x_i - \bar{x}_w$. The uncertainty of this deviation is calculated as a combination of the uncertainties of the result, $u(x_i)$, and the uncertainty of the weighted mean. In this case, the uncertainty of the weighted mean is taken as $u_{int}(\bar{x}_w)$. The uncertainty of the deviation from the weighted mean is given by equation (7), which includes a minus sign to take into account the correlation between the two uncertainties (it would be a plus sign if dealing with uncorrelated uncertainties, such as when comparing data from two separate laboratories).

$$u(x_i - \bar{x}_w) = \sqrt{[u(x_i)]^2 - [u_{int}(\bar{x}_w)]^2} \qquad (7)$$

Values for the weighted mean, internal and external standard deviations, deviation from weighted mean and its corresponding uncertainty and are calculated for each gauge block, and reported in Tables 10(a) through 10(d).

9.2 Analysis using E_n values

A check for statistical consistency of the results with their associated uncertainties can be made by calculating the E_n value for each laboratory, where E_n is defined as the ratio of the deviation from the weighted mean, divided by the uncertainty of this deviation:

$$E_n = \frac{x_i - \bar{x}_w}{\sqrt{[u(x_i)]^2 - [u_{int}(\bar{x}_w)]^2}} \qquad (8)$$

E_n values for each laboratory have been calculated and are also reported in Tables 10(a) through 10(d).

9.3 Birge ratio test

The statistical consistency of a comparison can also be investigated by the so-called Birge ratio R_B [17], which compares the observed spread of the results with the spread expected from the individual reported uncertainties.

The application of least squares algorithms and the χ^2-test leads to the Birge ratio:

$$R_B = \frac{u_{ext}(\bar{x}_w)}{u_{int}(\bar{x}_w)} \qquad (9)$$

The Birge ratio has an expectation value of R_B = 1, when considering standard uncertainties. For a coverage factor of k = 2, the expectation value is increased and the data in a comparison are consistent provided that

$$R_B < \sqrt{1 + \sqrt{8/(I-1)}} \qquad (10)$$

where I is the number of laboratories. For the case I = 12, a value of R_B < 1.36 indicates consistency (for k = 2).

9.4 Results of all participants

Tables 10(a) through 10(d) present the analysis of the results for the four gauge blocks, as described in sections 9.1 through 9.3, with displayed values rounded to the nearest nanometre.

175 mm S/N 6071						
LAB	x_i (µm)	$u_c(x_i)$ (µm)	w_i	$x_i - x_w$	$u(x_i - x_w)$	E_n
IMGC	0.140	0.028	0.048	-0.027	0.027	-0.97
PTB	0.122	0.013	0.223	-0.045	0.011	-3.89
NPL	0.161	0.030	0.042	-0.006	0.029	-0.19
NIST	0.142	0.016	0.147	-0.025	0.015	-1.66
INMETRO	0.150	0.020	0.094	-0.017	0.019	-0.87
NRC	0.125	0.027	0.052	-0.042	0.026	-1.58
NRLM	0.148	0.019	0.106	-0.018	0.018	-1.03
NIM	0.194	0.019	0.102	0.027	0.018	1.51
CSIRO	0.154	0.023	0.071	-0.013	0.022	-0.57
CSIR	0.180	0.110	0.003	0.013	0.110	0.12
SMU	0.410	0.038	0.026	0.243	0.038	6.49
VNIIM	0.312	0.021	0.085	0.145	0.020	7.24
Weighted mean, x_w	0.167					
C	3.770E-05					
$u_{int}(x_w)$	0.0061					
$u_{ext}(x_w)$	0.019					
R_B	3.168					

Table 10(a) Data analysis for measurements of the 175 mm gauge block, S/N 6071, showing deviations from weighted mean, and associated uncertainties. Also shown is the Birge ratio for this data set, R_B ,and the individual E_n values.

500 mm S/N 6071						
LAB	x_i (µm)	$u_c(x_i)$ (µm)	w_i	$x_i - x_w$	$u(x_i - x_w)$	E_n
IMGC	0.916	0.033	0.077	-0.013	0.032	-0.41
PTB	0.915	0.016	0.326	-0.014	0.013	-1.06
NPL	0.962	0.038	0.058	0.033	0.037	0.90
NIST	0.908	0.023	0.158	-0.021	0.021	-0.99
INMETRO	0.930	0.020	0.208	0.001	0.018	0.06
NRC	0.881	0.067	0.019	-0.048	0.066	-0.72
NRLM	0.938	0.039	0.054	0.009	0.038	0.24
NIM	1.007	0.060	0.023	0.078	0.060	1.31
CSIRO	0.885	0.050	0.033	-0.044	0.049	-0.89
CSIR	0.980	0.150	0.004	0.051	0.150	0.34
SMU	1.320	0.074	0.015	0.391	0.073	5.33
VNIIM	0.952	0.056	0.027	0.023	0.055	0.42
Weighted mean, x_w	0.929					
C	8.335E-05					
$u_{int}(x_w)$	0.0091					
$u_{ext}(x_w)$	0.016					
R_B	1.751					

Table 10(b) Data analysis for measurements of the 500 mm gauge block, S/N 6071, showing deviations from weighted mean, and associated uncertainties. Also shown is the Birge ratio for this data set, R_B ,and the individual E_n values.

500 mm S/N 3701						
LAB	x_i (µm)	$u_c(x_i)$ (µm)	w_i	$x_i - x_w$	$u(x_i - x_w)$	E_n
IMGC	0.814	0.033	0.073	-0.011	0.032	-0.34
PTB	0.807	0.015	0.353	-0.018	0.012	-1.47
NPL	0.861	0.038	0.055	0.036	0.037	0.98
NIST	0.781	0.023	0.150	-0.044	0.021	-2.06
INMETRO	0.830	0.020	0.199	0.005	0.018	0.30
NRC	0.786	0.066	0.018	-0.039	0.065	-0.59
NRLM	0.858	0.039	0.053	0.034	0.038	0.89
NIM	0.912	0.060	0.022	0.087	0.060	1.46
CSIRO	0.818	0.050	0.032	-0.007	0.049	-0.14
CSIR	0.870	0.150	0.004	0.045	0.150	0.30
SMU	1.250	0.072	0.015	0.425	0.071	5.95
VNIIM	0.868	0.056	0.025	0.043	0.055	0.78
Weighted mean, x_w	0.825					
C	7.953E-05					
$u_{int}(x_w)$	0.0089					
$u_{ext}(x_w)$	0.018					
R_B	2.020					

Table 10(c) *Data analysis for measurements of the 500 mm gauge block, S/N 3701, showing deviations from weighted mean, and associated uncertainties. Also shown is the Birge ratio for this data set, R_B ,and the individual E_n values.*

900 mm S/N 3701						
LAB	x_i (µm)	$u_c(x_i)$ (µm)	w_i	$x_i - x_w$	$u(x_i - x_w)$	E_n
IMGC	2.033	0.042	0.114	0.006	0.040	0.15
PTB	1.983	0.021	0.458	-0.044	0.015	-2.85
NPL	2.057	0.052	0.075	0.030	0.050	0.60
NIST	2.075	0.060	0.056	0.048	0.058	0.82
INMETRO	2.020	0.035	0.165	-0.007	0.032	-0.22
NRC	2.004	0.118	0.015	-0.023	0.117	-0.20
NRLM	2.070	0.068	0.044	0.042	0.066	0.64
NIM	2.160	0.136	0.011	0.133	0.135	0.98
CSIRO	1.982	0.087	0.027	-0.045	0.086	-0.53
CSIR	2.010	0.250	0.003	-0.017	0.250	-0.07
SMU	2.900	0.125	0.013	0.873	0.124	7.03
VNIIM	2.165	0.100	0.020	0.138	0.099	1.39
Weighted mean, x_w	2.027					
C	2.019E-04					
$u_{int}(x_w)$	0.0142					
$u_{ext}(x_w)$	0.033					
R_B	2.292					

Table 10(d) *Data analysis for measurements of the 900 mm gauge block, S/N 3701, showing deviations from weighted mean, and associated uncertainties. Also shown is the Birge ratio for this data set, R_B ,and the individual E_n values.*

All four analysis tables [10(a) through 10(d)] show a Birge ratio considerably greater than 1, indicating that the data is not consistent with the stated uncertainties. This may be due to the participants under-estimating their uncertainties, or by erroneous data being reported. Plotting the E_n values as a histogram, in Figure 5, allows identification of possible outliers, as well as giving a graphical view of the distribution of the data.

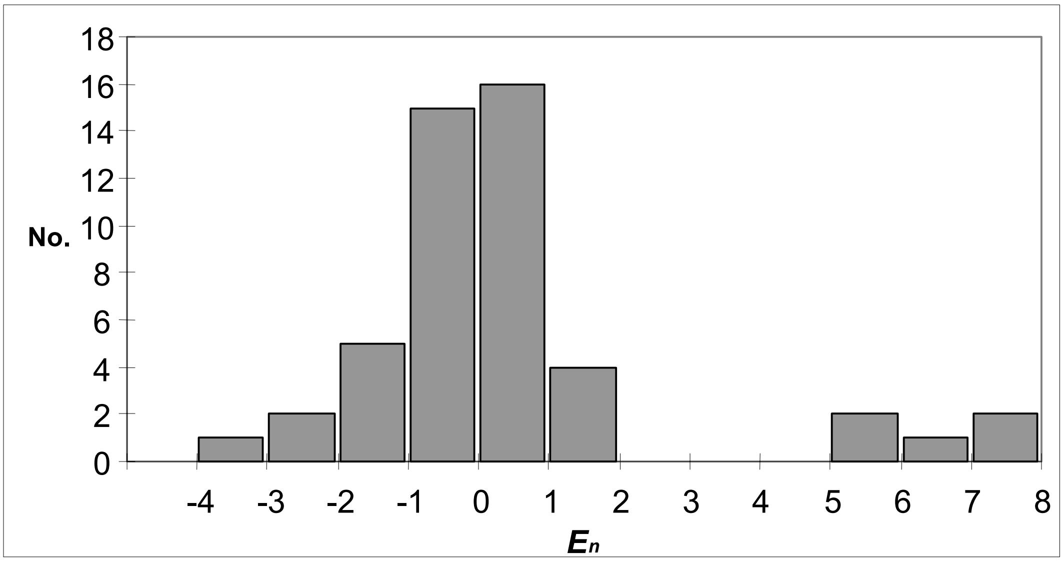


Figure 5 *Histogram of E_n values from the whole 48 element dataset of deviation from weighted mean, based on measurement results and standard uncertainties reported by the participants.*

As can be seen in Figure 5, the data is clustered into two groups, the larger group with E_n values from –3 to +2 and a smaller group with E_n values from +6 to +8. Using standard uncertainties in equation (8), one would expect 95% (45.6 out of 48) of the E_n values to be within the range –2 to +2, so the small group of 5 results on the right side of Figure 5 may be candidates for treatment as outliers. Examination of the data in Tables 10(a) through 10(d) shows that this small group of data corresponds to the results for all four gauges reported by laboratory SMU and the result for the 175 mm gauge block reported by VNIIM. The decision to exclude data as 'outliers' should not be based solely on statistical analysis, particularly where the set of data is relatively small, as is the case for this comparison. Thus SMU and VNIIM were contacted to discuss these findings.

SMU had already reported problems with temperature control at the time when they made their measurements and they had expressed concern that their data may not be as reliable as they would wish. Further analysis by SMU, after the results had been made public in the first draft of the Final Report, showed that the temperatures had been measured incorrectly and the real temperatures of the gauge blocks were higher than at first thought, and so the correction for the thermal expansion was incorrect. Thus the original data from SMU is identified as containing a significant bias, thus it is reasonable to treat these four results as outliers, such that they do not contribute to the weighted mean.

For information, the corrected results for SMU were as follows:

Serial Number	Nominal length (mm)	Deviation from nominal (μm)	Uncertainty (1σ) (nm)
6071	175	+ 0.285	38
6071	500	+ 0.960	74
3701	500	+ 0.900	72
3701	900	+ 2.260	125

Table 11 *Results from laboratory SMU corrected after discovery of temperature measurement error, after publication of the first draft of the Final Report.*

Because the guidelines for Key Comparisons do not allow a laboratory to change results after the data is made public, it is not possible to use the SMU corrected results in determining the reference value. Also despite adjustments following discovery of the error, the SMU results for the 175 mm and 900 mm gauge blocks still appear as outliers and this is confirmed by communication with SMU. Because the SMU data is now known to contain errors and is not representative of their standard measurement technique, it was decided to withdraw the data from the comparison.

One of the results in the group to the right of the histogram in Figure 5 corresponds to the VNIIM result for the 175 mm gauge block. Communication with VNIIM indicated that the measurement technique used for this gauge block was different to the techniques used for the other 3 gauge blocks. The measurement technique relies on the multiplication of an etalon distance of 100 mm, to reach the size of the gauge block to be measured. Because the nominal size of 175 mm was not an integer multiple of 100 mm, an additional gauge block was needed of size 25 mm. The uncertainties associated with the use of this gauge block were unknown and the overall uncertainty of measurement was probably underestimated. VNIIM agreed that it was likely that this measurement point was subject to additional uncertainties and should be treated as an outlier. Therefore it does not contribute to the weighted mean but since it is a valid result, it remains in the final table of results.

9.5 Analysis of results, outliers excluded from weighted mean

As discussed in section 9.4, five results were identified as possible outliers and discussion with the laboratories concerned has confirmed this analysis. It is likely that these outlying results are biasing the weighted mean values and contributing to uncertainties which cause the Birge ratios and the E_n values to be greater than expected.

Therefore, in order to progress with more accurate analysis of the data, the analysis of section 9.4 is repeated, but with the five outlying data points excluded from the determination of the weighted mean and the SMU results withdrawn. The results of this analysis are presented in Table 12(a) through Table 12(d) and in Figure 6. Displayed results are rounded to the nearest nanometre.

175 mm S/N 6071						
LAB	x_i (μm)	$u_c(x_i)$ (μm)	w_i	$x_i - x_w$	$u(x_i - x_w)$	E_n
IMGC	0.140	0.028	0.054	-0.005	0.027	-0.20
PTB	0.122	0.013	0.251	-0.023	0.011	-2.08
NPL	0.161	0.030	0.047	0.016	0.029	0.53
NIST	0.142	0.016	0.166	-0.003	0.015	-0.24
INMETRO	0.150	0.020	0.106	0.005	0.019	0.24
NRC	0.125	0.027	0.058	-0.020	0.026	-0.78
NRLM	0.148	0.019	0.119	0.003	0.018	0.16
NIM	0.194	0.019	0.115	0.049	0.018	2.69
CSIRO	0.154	0.023	0.080	0.009	0.022	0.39
CSIR	0.180	0.110	0.004	0.035	0.110	0.31
VNIIM	0.312	0.021	N/A	0.167	0.020	8.34
Weighted mean, x_w	0.145					
C	4.243E-05					
$u_{int}(x_w)$	0.0065					
$u_{ext}(x_w)$	0.007					
R_B	1.100					

Table 12(a) Data analysis for measurements of the 175 mm gauge block, S/N 6071, with data from one participant excluded from the weighted mean as an outlier and one result withdrawn, showing deviations from weighted mean, and associated uncertainties. Also shown is the Birge ratio for this data set, R_B ,and the individual E_n values.

500 mm S/N 6071						
LAB	x_i (µm)	$u_c(x_i)$ (µm)	w_i	$x_i - x_w$	$u(x_i - x_w)$	E_n
IMGC	0.916	0.033	0.078	-0.007	0.032	-0.22
PTB	0.915	0.016	0.331	-0.008	0.013	-0.60
NPL	0.962	0.038	0.059	0.039	0.037	1.06
NIST	0.908	0.023	0.160	-0.015	0.021	-0.71
INMETRO	0.930	0.020	0.212	0.007	0.018	0.40
NRC	0.881	0.067	0.019	-0.042	0.066	-0.63
NRLM	0.938	0.039	0.055	0.015	0.038	0.40
NIM	1.007	0.060	0.023	0.084	0.060	1.41
CSIRO	0.885	0.050	0.034	-0.038	0.049	-0.77
CSIR	0.980	0.150	0.004	0.057	0.150	0.38
VNIIM	0.952	0.056	0.027	0.029	0.055	0.53
Weighted mean, x_w	0.923					
C	8.464E-05					
$u_{int}(x_w)$	0.0092					
$u_{ext}(x_w)$	0.007					
R_B	0.732					

Table 12(b) Data analysis for measurements of the 500 mm gauge block, S/N 6071, with data from one participant withdrawn as an outlier, showing deviations from weighted mean, and associated uncertainties. Also shown is the Birge ratio for this data set, R_B ,and the individual E_n values.

500 mm S/N 3701						
LAB	x_i (µm)	$u_c(x_i)$ (µm)	w_i	$x_i - x_w$	$u(x_i - x_w)$	E_n
IMGC	0.814	0.033	0.074	-0.004	0.032	-0.13
PTB	0.807	0.015	0.359	-0.011	0.012	-0.92
NPL	0.861	0.038	0.056	0.043	0.037	1.16
NIST	0.781	0.023	0.153	-0.037	0.021	-1.75
INMETRO	0.830	0.020	0.202	0.012	0.018	0.67
NRC	0.786	0.066	0.019	-0.032	0.065	-0.49
NRLM	0.858	0.039	0.054	0.040	0.038	1.07
NIM	0.912	0.060	0.022	0.094	0.060	1.57
CSIRO	0.818	0.050	0.032	0.000	0.049	0.00
CSIR	0.870	0.150	0.004	0.052	0.150	0.35
VNIIM	0.868	0.056	0.026	0.050	0.055	0.90
Weighted mean, x_w	0.818					
C	8.077E-05					
$u_{int}(x_w)$	0.0090					
$u_{ext}(x_w)$	0.009					
R_B	0.972					

Table 12(c) Data analysis for measurements of the 500 mm gauge block, S/N 3701, with data from one participant withdrawn as an outlier, showing deviations from weighted mean, and associated uncertainties. Also shown is the Birge ratio for this data set, R_B ,and the individual E_n values.

900 mm S/N 3701						
LAB	x_i (µm)	$u_c(x_i)$ (µm)	w_i	$x_i - x_w$	$u(x_i - x_w)$	E_n
IMGC	2.033	0.042	0.116	0.017	0.039	0.44
PTB	1.983	0.021	0.464	-0.033	0.015	-2.13
NPL	2.057	0.052	0.076	0.041	0.050	0.83
NIST	2.075	0.060	0.057	0.059	0.058	1.02
INMETRO	2.020	0.035	0.167	0.004	0.032	0.14
NRC	2.004	0.118	0.015	-0.012	0.117	-0.10
NRLM	2.070	0.068	0.044	0.054	0.066	0.81
NIM	2.160	0.136	0.011	0.144	0.135	1.07
CSIRO	1.982	0.087	0.027	-0.034	0.086	-0.39
CSIR	2.010	0.250	0.003	-0.006	0.250	-0.02
VNIIM	2.165	0.100	0.020	0.149	0.099	1.51
Weighted mean, x_w	2.016					
C	2.045E-04					
$u_{int}(x_w)$	0.0143					
$u_{ext}(x_w)$	0.013					
R_B	0.915					

Table 12(d) Data analysis for measurements of the 900 mm gauge block, S/N 3701, with data from one participant withdrawn as an outlier, showing deviations from weighted mean, and associated uncertainties. Also shown is the Birge ratio for this data set, R_B ,and the individual E_n values.

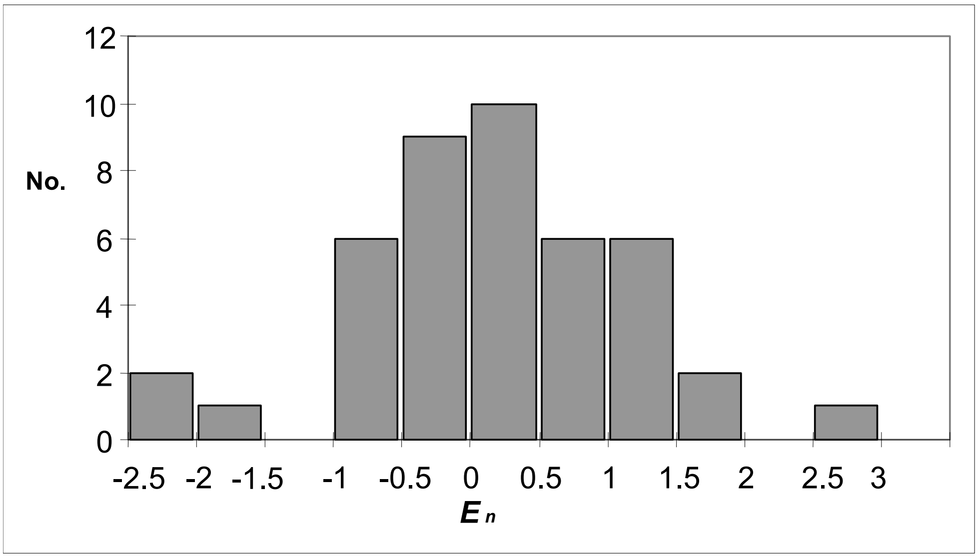


Figure 6 Histogram of E_n values from the dataset of deviation from weighted mean, based on measurement results and standard uncertainties reported by the participants, with exclusion of five outlier results (43 results remaining).

Examination of Tables 12(a) though 12(d) shows that the Birge ratios for the four sets of data have all decreased and are much closer to unity with the possible exception of that of the 500 mm gauge block, S/N 6071 which is now much lower than unity. It should be noted that the Birge ratios for the two 500 mm gauge blocks are somewhat different, whereas one would expect similarity due to similar measurement procedures and uncertainties for the two gauge blocks. The differences may relate to physical properties of the two gauge blocks being different, although there is little difference in the properties of these gauge blocks as determined by the pilot laboratory (thermal expansion coefficient, flatness, variation in length, central length stability).

Based on the Birge ratios, it appears that overall, the uncertainty estimates are now in good agreement with the observed spread in data, with a possible over-estimation of the uncertainty for the measurements of the 500 mm S/N 6071 gauge block.

Examination of Figure 6 indicates that for 40 of the 43 results, *i.e.* 93%, the E_n ratio has a magnitude less than or equal to 2. This compares favourably with the expectation of 95% of the results being within the stated uncertainties, at k = 2.

A summary of all of the (non-outlier) measurement data is represented in Figure 7.

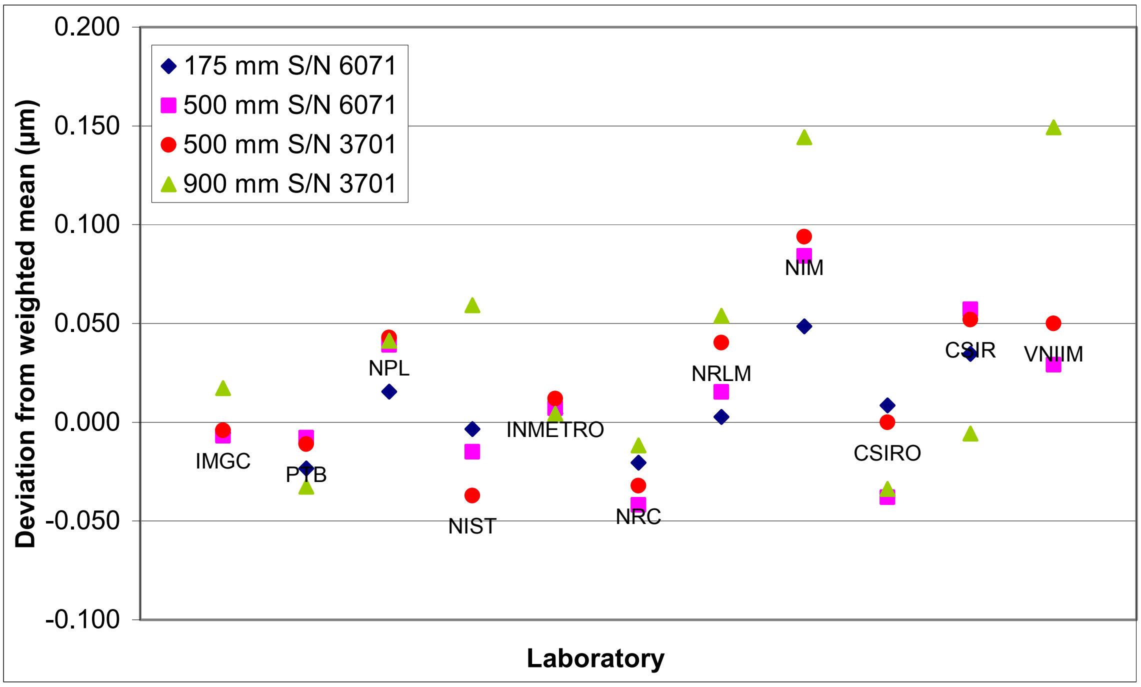


Figure 7 Laboratory deviations from weighted means, outliers excluded.

9.6 Discussion of results

Excluding outliers, the total spread of the whole set of results, over the four gauge blocks is 200 nm, with most of the participants reporting results which agree within ± 50 nm, which represents good agreement for relatively long dimensional artefact calibrations. In some part, this may be due to the availability of pre-determined thermal expansion coefficients for the artefacts. For example, the 900 mm gauge block had a nominal expansion coefficient (according to ISO 3650) of 11.5 ± 0.5 x 10^{-6} K^{-1} whereas its determined value was 10.824 x 10^{-6} K^{-1}, a difference of 0.676 x 10^{-6} K^{-1}. For the laboratories which reported the largest temperature deviation from 20 °C of up to 0.104 °C, this difference could have contributed an additional error of 63 nm. Future comparisons should take this into account, and pre-determined values of parameters such as thermal expansion coefficient should not be routinely available, if the purpose of the key comparison is to test participants' ability to measure 'unknown' artefacts from customers.

With regard to Figure 7, there may be evidence for systematic deviations of several laboratories, with respect to each other. Whilst the variations in the results for each particular laboratory are generally within the specified uncertainties (which are difficult to show clearly in the graph), there are some possible trends which can be observed. For example, laboratories NPL, NRLM, NIM, CSIR and VNIIM generally measure longer, whereas laboratories PTB, NIST, NRC, and CSIRO generally measure shorter. These differences may be due to factors such as phase correction determination or compensation of the support position, vertical rather than horizontal measurement, refractive index determination or wringing effects. These are now examined.

9.6.1 Phase correction

Although the specification standard ISO 3650 requires that "Corrections shall be made to the calculations for significant influences; *e.g.* : … surface texture and optical phase changes on the reflection of the light wave", *i.e.* the so called 'phase correction', only three participants (PTB, NRLM, NIM) performed this correction. Other participants were not able to perform experiments to determine this correction and, instead, made no correction but expanded their uncertainty to take this into account. These laboratories reported that this was the usual procedure at their institutes when offering this measurement service for customers. Of the three laboratories which reported a phase correction, two participants used steel platens (PTB, NRLM), the other used glass (NIM), and the phase correction values ranged from –6 nm to

+60 nm, respectively. These three laboratories are on opposite sides (positive, negative) of the reference value line in Figure 7, however both laboratories which performed a phase correction and made measurements using steel platens, determined a negative phase correction, of small magnitude, indicating close similarity between gauge block and platen materials. In a recent key comparison for short gauge blocks, CCL-K1, the majority of the participants did determine a phase correction, which, for steel gauge blocks measured on steel platens, ranged from -41 nm to +7 nm. In comparison CCL-K2, it seems reasonable to expect that the undetermined phase correction values for measurements using steel platens would cover a similar range, if the same types of steel platens were used. This seems to be reflected in the attribution of uncertainties for the phase correction in the range 6 nm to 40 nm (k = 2), by participants which did not perform a phase correction determination. Figure 6 also shows that the mean of the E_n values is not zero, but is quite positive, indicating a possible bias in the weighted mean values.

9.6.2 Vertical or horizontal measurement

The three laboratories which made (some) measurements with the gauges standing vertically are distributed both above and below the reference value line, however this is some disagreement on the magnitude of the correction terms applied.

9.6.3 Compensation of platen weight

Two participants reported that the weight of the platen was compensated by a counterbalance system, whereas other participants moved the supports instead. Both sets of participants are equally distributed above and below the reference line.

9.6.4 Compensation for refractive index

There is no observed correlation between reported result and the method used to determine the refractive index correction.

9.7 Further discussion

The fact that the majority of participants did not determine a phase correction causes a significant problem. The *Guide to the Expression of Uncertainty in Measurement* requires that the result of a measurement has to be corrected for all recognised significant systematic effects. In the case of key comparison CCL-K2, there is the possibility that the undetermined phase correction values are significant, when compared with the deviations from the weighted mean or when compared with other uncertainty sources. Normally, such results would be excluded from the calculation of the mean, and the laboratories concerned would have to accept the possibility of their deviation from the mean value being significantly influenced by their (undetermined) phase correction. However, in the case of CCL-K2, this would lead to only three laboratories contributing to the mean value and the difference between the results of these three laboratories is rather large.

Although the participants which did not determine a phase correction have correspondingly increased the uncertainty contribution of this influence parameter, the analysis using the weighted mean is still not fair for all participants. The majority of the undetermined phase corrections are likely to be negative due to normal roughnesses of gauge block and platen surfaces, as mentioned in section 9.6.1. Therefore participants which did not determine a phase correction are probably biasing the weighted means, towards higher values. The increased uncertainties for these participants will help maintain their E_n values close to 1, but the other laboratories will have their deviations from the mean values altered, but with no commensurate change to their uncertainty to offset this. Thus the E_n values of laboratories which strictly followed the protocol may be unduly influenced by those laboratories which were unable to make all the necessary corrections. In effect, the weighted mean values have been influenced by the selection of the laboratories which were chosen to be participants.

Figure 7 shows information about the measurements of all four gauge blocks, linked through the determined weighted means, allowing some analysis of length dependent and length independent errors. The fact that the coefficient of thermal expansion was given for all four gauge blocks means one of the largest length dependent sources of error has been significantly reduced. Other potentially large length dependent error sources are also well controlled: use of standard and well known equations for refractive index determination, stabilized laser or Kr lamp sources for measurement wavelengths. This is reflected in the clustering of data in Figure 7, where deviations from weighted mean values are, for many participants, grouped in clusters spanning approximately 50 nm, or less. However the centres of the clusters differ from each other by more than 50 nm, in many cases, indicating uncontrolled length independent errors. There may also be indications of uncontrolled length dependent errors for one or two participants.

9.8 Normalised differences between laboratories

Because there is some question concerning the correctness of choosing a reference value when it is suspected that there may be non-symmetrically distributed uncorrected biases in some results, an alternative analysis is to examine only normalised differences between laboratories' results. The normalised difference takes into account the fact that the uncertainties of the two laboratories are uncorrelated and therefore the normalised difference between laboratories i and j, is given by

$$\hat{d}_{ij} = \frac{x_i - x_j}{\sqrt{[u(x_i)]^2 + [u(x_j)]^2}} \qquad (11)$$

These values are given, for information, in Tables 13(a) though 13(d). A value with magnitude less than 2 indicates the differences are commensurate with the stated uncertainties, for a *k* = 2 coverage factor.

175 mm S/N 6071											
	IMGC	**PTB**	**NPL**	**NIST**	**INMETRO**	**NRC**	**NRLM**	**NIM**	**CSIRO**	**CSIR**	**VNIIM**
IMGC	0.00	-0.58	0.51	0.06	0.29	-0.39	0.24	1.59	0.39	0.35	4.91
PTB	0.58	0.00	1.19	0.97	1.17	0.10	1.14	3.11	1.21	0.52	7.69
NPL	-0.51	-1.19	0.00	-0.56	-0.31	-0.89	-0.36	0.93	-0.19	0.17	4.12
NIST	-0.06	-0.97	0.56	0.00	0.31	-0.54	0.25	2.08	0.43	0.34	6.44
INMETRO	-0.29	-1.17	0.31	-0.31	0.00	-0.74	-0.07	1.59	0.13	0.27	5.59
NRC	0.39	-0.10	0.89	0.54	0.74	0.00	0.70	2.08	0.82	0.49	5.47
NRLM	-0.24	-1.14	0.36	-0.25	0.07	-0.70	0.00	1.70	0.19	0.28	5.80
NIM	-1.59	-3.11	-0.93	-2.08	-1.59	-2.08	-1.70	0.00	-1.34	-0.13	4.15
CSIRO	-0.39	-1.21	0.19	-0.43	-0.13	-0.82	-0.19	1.34	0.00	0.23	5.07
CSIR	-0.35	-0.52	-0.17	-0.34	-0.27	-0.49	-0.28	0.13	-0.23	0.00	1.18
VNIIM	-4.91	-7.69	-4.12	-6.44	-5.59	-5.47	-5.80	-4.15	-5.07	-1.18	0.00

500 mm S/N 6071											
	IMGC	**PTB**	**NPL**	**NIST**	**INMETRO**	**NRC**	**NRLM**	**NIM**	**CSIRO**	**CSIR**	**VNIIM**
IMGC	0.00	-0.03	0.91	-0.20	0.36	-0.47	0.43	1.32	-0.52	0.42	0.55
PTB	0.03	0.00	1.14	-0.25	0.59	-0.49	0.55	1.47	-0.57	0.43	0.64
NPL	-0.91	-1.14	0.00	-1.22	-0.75	-1.05	-0.44	0.63	-1.23	0.12	-0.15
NIST	0.20	0.25	1.22	0.00	0.72	-0.38	0.66	1.53	-0.42	0.47	0.73
INMETRO	-0.36	-0.59	0.75	-0.72	0.00	-0.70	0.19	1.21	-0.84	0.33	0.37
NRC	0.47	0.49	1.05	0.38	0.70	0.00	0.74	1.40	0.05	0.60	0.81
NRLM	-0.43	-0.55	0.44	-0.66	-0.19	-0.74	0.00	0.95	-0.84	0.27	0.20
NIM	-1.32	-1.47	-0.63	-1.53	-1.21	-1.40	-0.95	0.00	-1.56	-0.17	-0.67
CSIRO	0.52	0.57	1.23	0.42	0.84	-0.05	0.84	1.56	0.00	0.60	0.89
CSIR	-0.42	-0.43	-0.12	-0.47	-0.33	-0.60	-0.27	0.17	-0.60	0.00	-0.17
VNIIM	-0.55	-0.64	0.15	-0.73	-0.37	-0.81	-0.20	0.67	-0.89	0.17	0.00

500 mm S/N 3701											
	IMGC	**PTB**	**NPL**	**NIST**	**INMETRO**	**NRC**	**NRLM**	**NIM**	**CSIRO**	**CSIR**	**VNIIM**
IMGC	0.00	-0.19	0.93	-0.82	0.41	-0.38	0.87	1.42	0.07	0.36	0.83
PTB	0.19	0.00	1.32	-0.95	0.92	-0.31	1.24	1.69	0.21	0.42	1.05
NPL	-0.93	-1.32	0.00	-1.80	-0.72	-0.98	-0.05	0.71	-0.68	0.06	0.10
NIST	0.82	0.95	1.80	0.00	1.61	0.07	1.72	2.03	0.67	0.59	1.44
INMETRO	-0.41	-0.92	0.72	-1.61	0.00	-0.64	0.65	1.29	-0.22	0.26	0.64
NRC	0.38	0.31	0.98	-0.07	0.64	0.00	0.95	1.41	0.39	0.51	0.95
NRLM	-0.87	-1.24	0.05	-1.72	-0.65	-0.95	0.00	0.75	-0.64	0.07	0.14
NIM	-1.42	-1.69	-0.71	-2.03	-1.29	-1.41	-0.75	0.00	-1.20	-0.26	-0.53
CSIRO	-0.07	-0.21	0.68	-0.67	0.22	-0.39	0.64	1.20	0.00	0.33	0.67
CSIR	-0.36	-0.42	-0.06	-0.59	-0.26	-0.51	-0.07	0.26	-0.33	0.00	-0.01
VNIIM	-0.83	-1.05	-0.10	-1.44	-0.64	-0.95	-0.14	0.53	-0.67	0.01	0.00

900 mm S/N 3701											
	IMGC	**PTB**	**NPL**	**NIST**	**INMETRO**	**NRC**	**NRLM**	**NIM**	**CSIRO**	**CSIR**	**VNIIM**
IMGC	0.00	-1.06	0.36	0.57	-0.24	-0.23	0.46	0.89	-0.53	-0.09	1.22
PTB	1.06	0.00	1.32	1.45	0.91	0.18	1.22	1.29	-0.01	0.11	1.78
NPL	-0.36	-1.32	0.00	0.23	-0.59	-0.41	0.15	0.71	-0.74	-0.18	0.96
NIST	-0.57	-1.45	-0.23	0.00	-0.79	-0.54	-0.06	0.57	-0.88	-0.25	0.77
INMETRO	0.24	-0.91	0.59	0.79	0.00	-0.13	0.65	1.00	-0.41	-0.04	1.37
NRC	0.23	-0.18	0.41	0.54	0.13	0.00	0.48	0.87	-0.15	0.02	1.04
NRLM	-0.46	-1.22	-0.15	0.06	-0.65	-0.48	0.00	0.60	-0.79	-0.23	0.79
NIM	-0.89	-1.29	-0.71	-0.57	-1.00	-0.87	-0.60	0.00	-1.10	-0.53	0.03
CSIRO	0.53	0.01	0.74	0.88	0.41	0.15	0.79	1.10	0.00	0.11	1.38
CSIR	0.09	-0.11	0.18	0.25	0.04	-0.02	0.23	0.53	-0.11	0.00	0.58
VNIIM	-1.22	-1.78	-0.96	-0.77	-1.37	-1.04	-0.79	-0.03	-1.38	-0.58	0.00

Tables 13(a) to 13(d) Normalised differences between laboratories' results

10 Conclusions

From the CCL-K2 long gauge block key comparison, the following conclusions can be drawn:

- It took two years from the decision to carry out this comparison until the protocol document was finally issued. This reflected the degree of discussion necessary to agree on the technical basis of the comparison, the likely timetable, list of participants, and to reach agreement on the issue of hand carriage of the gauge blocks during transportation.

- From the start of the comparison, the time taken to perform the measurements, including transportation, was almost as planned, namely 23 months (18 planned). This represents a typical length of time required for a comparison of physical artefacts in dimensional metrology and was only achieved by careful planning and co-operation of the participants. The fact that flights had to be booked in advance (for the hand carriage) may have helped participants focus on keeping to the timetable.

- The timetable was adjusted for two laboratories at the end of the comparison, which reported temperature control problems to the pilot laboratory prior to circulation in their region. Approval from participants allowed a delay to be accepted. However, further delays were encountered due to customs regulations in one country.

- The decision to limit the number of participating laboratories to about a dozen was very reasonable. The surface quality of the gauge blocks at the end of the comparison would not have allowed for many additional measurements without the risk of seriously degrading the measurement results.

- The decision to limit the transportation to hand-carriage was vindicated by the lack of damage to the gauge blocks during the comparison. A previous EUROMET comparison (EUROMET Project 254 [18]) showed that long gauge blocks are particularly vulnerable to damage during transportation. Fortunately, the airlines used during this comparison were co-operative in allowing the gauge block case into the cabin during the flights.

- The close agreement of many results may be due to the inclusion in the protocol document of pre-determined values for the linear coefficient of thermal expansion for the supplied gauge blocks. This allowed laboratories to make accurate corrections for any measurements performed at temperatures away from the standard temperature of 20 °C. Whilst this helps ensure a uniformity of results, it does not reflect a typical measurement for a customer, where the expansion coefficient is not known by the laboratory other than the nominal value attributed to the gauge block material. Thus the results of this comparison may represent an over-optimistic view of the mutual equivalence of the services offered by the participants. At the September 2000 meeting of the CCL-WGDM it was decided that in any future comparisons, no 'additional' data such as thermal expansion coefficients would be supplied. Only data that would normally be available to a customer or calibration laboratory would be included in the protocol document. However the pilot laboratory would perform a check that the artefacts complied with any specification standards, in respect of any such data.

- The reliability of the weighted mean, or other estimators of the true value of the measurands has been discussed. Non-symmetrically distributed, uncorrected biases to some results (phase corrections) are thought to be possible explanations for some laboratories' differences from each other. When examining deviation of results from weighted mean results, this fact should be taken into account. The use of bilateral comparisons between two laboratories' results is expected to be more robust, as possible biases of the reference value (weighted mean) have no influence.

- The principal aim of this key comparison has been to determine the degree to which results of measurement of long gauge blocks made by a selection of NMIs can be deemed to be 'equivalent. This has been tested by measurement of four almost-unknown long gauge blocks, using techniques normally used by each participant for such measurements. This has resulted in a set of data which can be used by the metrology community to gain insight into degrees of equivalence of NMI measurements of long gauge blocks. However one should also try to maximise the scientific value of this comparison, It would be useful for each participant to examine their results and measurement processes in light of this report, and seek explanations for any significant offsets of their results from those of other laboratories.

11 Acknowledgements

The pilot laboratory would like to acknowledge the kind assistance, and willingness to fund the necessary travel, of all the colleagues in the participating laboratories for helping this comparison to run so smoothly. Also beneficial to the successful completion of the comparison was the assistance afforded by all the airlines which were kind enough to allow hand carriage of the gauge case in their aircraft.

Particular thanks are also due to PTB for kindly donating the long gauge blocks that were used in this comparison.

The assistance of colleagues Jennifer Decker (NRC) and Nick Brown (CSIRO) was welcome during the preparation of the protocol document. The protocol document and report of key comparison CCL-K1, written by Ruedi Thalmann (METAS) were also extremely useful for the purposes of preparing the corresponding documents for comparison CCL-K2.

The pilot laboratory would like to acknowledge the financial support for the pilot laboratory activities associated with this key comparison, which were funded from a Technology Transfer project of the UK Government's National Measurement System Programme for Length, 1999-2002, MPU 8/61.

12 References

[1] T. J. Quinn, Guidelines for key comparisons carried out by Consultative Committees, draft of 21. November 1997, BIPM, Paris.

[2] R. Thalmann, Swiss Federal Office of Metrology & Accreditation (METAS), Technical Protocol document, and Report Draft B, key comparison CCL-K1 (2001).

[3] J. E. Decker and J.R. Pekelsky, Uncertainty Evaluation for the Measurement of Gauge Blocks by Optical Interferometry, *Metrologia* **34,** 479-493 (1997).

[4] H. Darnedde, High-precision calibration of long gauge blocks using the vacuum wavelength comparator, *Metrologia,* **29**, 349-359 (1992).

[5] A. J. Lewis, Measurement of length, surface form and thermal expansion coefficient of length bars up to 1.5 m using multiple-wavelength phase-stepping interferometry, *Meas. Sci. Technol.*, **5**, 694-703 (1994).

[6] J. Beers, Measuring long gauge blocks with the NIST line scale interferometer, NIST Technical Note 1410 (1995).

[7] J. E. Decker & J. R. Pekelsky, Gauge block calibration by optical interferometry at the NRC Canada, NRC Document 40002 (1997).

[8] J. E. Decker, K. Bustraan, S. de Bonth, J. R. Pekelsky, Updates to the NRC gauge block interferometer, NRC Document 42753 (2000).

[9] B. Edlen, The refractive index of air, *Metrologia*, **2**, 71-80 (1966).

[10] K. P. Birch & M. J. Downs, Correction to the updated Edlén equation for the refractive index of air, *Metrologia*, **31**, 315-316 (1994).

[11] P. E. Ciddor, Refractive index of air: new equations for the visible and near infrared, *Appl. Opt.*, **35**, 1566 –1573 (1996).

[12] F. Beyer-Helms, Biegung von Endmaßen bei horizontaler Lagerung auf Scheiden, *PTB Mitteilungen*, 25-30 and 124-130 (1967).

[13] G. Bönsch, Interferometric calibration of an integrating sphere for determination of the roughness correction of gauge blocks, *Proc. SPIE*, **3477**, 152-160 (1998).

[14] G. Bönsch, Automatic gauge block measurements by phase stepping interferometry with three laser wavelengths, *Proc. SPIE*, **4401**, 1-10 (2001).

[15] D. C. Williams, The parallelism of a length bar with an end load, *J. Sci. Instrum.*, **39**, 608-610 (1962).

[16] Communication to participants from B. Frenkel via N. Brown, during discussion of CCL-K1 final report.

[17] *Statistical Analysis of Interlaboratory Comparisons*, EUROMET workshop held at NPL on 11.-12. November 1999, http://www.npl.co.uk/ssfm/download/documents/sss_m_00_173.pdf

[18] H. Darnedde and J. Helmcke, European comparison of long gauge blocks, *Metrologia* **33**, 485-491 (1996).

Appendix 1: Determination of the Key Comparison Reference Values

12.1 KCRVs and their uncertainties

In order to satisfy the requirements of the Mutual Recognition Arrangement, the so-called 'Key Comparison Reference Values' have been evaluated according to the method described in section 9 of the main document, *i.e.* the weighted mean is determined and the deviations from the weighted mean are then calculated. Data identified as outliers are excluded from the determination of the weighted mean. This method requires that the individual uncertainties from the laboratories were estimated according to a common approach (which should be the case, since all participants were requested to estimate the uncertainties according to the *ISO Guide*). If this is not the case, a single "wrong" value with a strongly underestimated (too small) uncertainty could strongly influence or even fully determine the weighted mean. On the other hand, a high quality measurement with overestimated uncertainty would contribute only a small amount to the reference value. The uncertainty of the weighted mean is based on the internal standard deviation of the contributing results.

Note that the results of the pilot laboratory contribute only once to the calculation of the reference values, namely its official measurement result. This excludes the preliminary, interim and final 'stability' measurements of the gauge blocks, performed by the pilot laboratory.

Serial number	Nominal length (mm)	Reference Value (nm)	Uncertainty in Reference Value (nm)
6071	175	+145	7
6071	500	+923	9
3701	500	+818	9
3701	900	+2 016	14

Table 14 Key Comparison Reference Values and associated standard uncertainties (in nm). The internal standard deviation is used for determination of the uncertainty The Key Comparison Reference Values represent the measured deviation from nominal size for the four gauge blocks used in the comparison.

It should be noted that there are several arguments presented in the main text of this report that indicate that the Key Comparison Reference Values may not be totally reliable estimators of the true values of each measurand, due to unknown, non-symmetric bias of some results. Even if such biases could be determined and corrected for, the Key Comparison Reference Values would have no significance in terms of the SI other than as the best estimates of the sizes of the four gauge blocks which were used in this comparison.

12.2 Artefact uncertainties

When calculating the degree of equivalence for each participant, it is necessary to consider additional sources of uncertainty, other than the uncertainty in the KCRV and the participant's uncertainty. Due to correlations between the participants' results and the weighted mean, the (k = 2) uncertainty of the difference from the weighted mean is usually given by

$$U(\Delta l) = 2\sqrt{u^2(x_i) - u^2(\bar{x}_{ref})} \quad (8)$$

where $u(x_i)$ and $u(\bar{x}_{ref})$ are the standard uncertainties of the laboratory result x_i and the reference value $\bar{x}_{ref}$. Although this is statistically correct, it fails to take into account the uncertainties associated with the artefacts, such as stability, accumulation of damage, *etc*. When performing calibrations for customers, these artefact-based errors are not included in any uncertainty analysis as the calibrations only determine 'on the day' values' and no allowance is made for subsequent drift, damage or misuse. However in a comparison such as CCL-K2 with an extended timescale, and measurements by a relatively large number of laboratories, artefact stability and the effects of damage accumulated during the circulation must be taken into account when comparing participants' results with the reference values.

This report therefore proposes the use of a third uncertainty component, $u(\bar{x}_{artefact})$, which is estimated from artefact performance in the comparison and other expert knowledge, and is therefore not correlated with the two other uncertainty components.

Possible contributions to the artefact uncertainty are:

- secular change in gauge block central length during the comparison;
- change in gauge block geometry (flatness, variation in length);
- accumulation of damage to wringing surfaces of gauge blocks affecting the wringing property;
- constraints imposed by the comparison (stabilisation timescales etc);
- change in phase correction due to surface wear;
- additional uncertainties, not normally considered significant enough to compensate for (*e.g.* pressure differentials between laboratories).

12.2.1 Secular change in length of gauge block

It is known that gauge blocks can exhibit a change in length, even if left undisturbed (see for example F. H. Rolt, *Gauges and Fine Measurements*, Macmillan, London, (1929), Chapter 10). Although the gauge blocks were specially selected as having a long history of stability, the artefacts of CCL-K2 have been transported around the world, subject to different temperatures and pressures (in transit) and to mechanical vibration. It is therefore reasonable to expect some change in length of the gauge blocks due to stress relief in the bulk material. The best estimator of this change in length is through the measurements of the pilot laboratory, before, during and after the circulation of the artefacts. The measured changes in length of the four gauge blocks are give in Table 15. It is difficult to estimate the uncertainty of this measurement as it depends on the wringing properties of the surfaces, but the best estimate standard uncertainty is the quadrature sum of ± 7 nm and ± $2 \times 10^{-8} L$. This is based on the pilot laboratory's measurement uncertainty, taking account of uncertainties which are common to all four measurements.

Serial number	Nominal length (mm)	Change in length (nm)	Standard uncertainty (nm)
6071	175	22	8
6071	500	40	12
3701	500	27	12
3701	900	38	20

Table 15 Measured changes in central length, as determined by the pilot laboratory.

For the purposes of determining the artefact uncertainty, this contribution will be treated as a rectangular distribution, of half width 11, 20, 14 and 19 nm, for the four gauge blocks, resulting in standard uncertainties of 7, 12, 9 and 11 nm, respectively.

12.2.2 Change in gauge block geometry

The changes in central length in section 12.2.1 already include the effect of any changes in surface geometry on central length. Of the two measured parameters, the change in flatness is likely to have the larger effect as it affects the wringing quality. Assuming that the laboratories were able to measure at the centre of the face, those that used a wrung platen will have an error caused by imperfect wringing at the wrung face of the gauge block. Previous work (G. Bönsch, *Proc. SPIE*, **3477**, 199-210 (1998)) has shown a typical error of 12 nm in the central length measurement of a gauge block caused by a flatness error of approximately 70 nm. The largest change in flatness of the four CCL-K2 gauge blocks was 63 nm for the 900 mm gauge block, which was already displaying a flatness error of ~100 nm. Therefore one could estimate that the change in flatness could contribute an additional length measurement error of the order of 12 nm. This value is well within the length changes detailed in section 12.2.1 and so no additional error term is proposed.

12.2.3 Damage to wringing surfaces

During the comparison, the surface quality was observed to become degraded due to repeated wringing causing scratching of the surface. No attempt was made to re-lap the surfaces, so it must be assumed that the wringing quality changed during the comparison. This was also reported by some participants. However, the pilot laboratory made measurements at the start and end of the comparison, using the same platens, so any change in length due to changes in surface quality should already be present in the data in section 12.2.2, so no further uncertainty component is proposed.

12.2.4 Constraints imposed by protocol

The only change from normal operating procedure required by the protocol document was for completion of all actions, including unpacking, preparation, measurement, packing and onwards transportation, within a 1 month timescale. This may have led to some participants using a shorter stabilisation time than usual. Work by Decker et al (J. E. Decker, *Metrologia,* **38**, 269-272 (2001)) has indicated the possibility of temporary length instability at the 1×10^{-8} L level for a period of up to 5 days after transportation. However, after this period, the length of the gauge block was stable. No information is available concerning stabilisation times used by participants, however the need to make two measurements (different wrings) would indicate that the second measurement at least would be performed with the gauge block in a stable configuration. Also, it is anticipated that participants would have made several measurements of the gauge blocks, and averaged the results. Therefore it is unlikely that this uncertainty component is any greater than ~5×10^{-9} L. Taking this as the full width of a rectangular distribution, leads to a standard uncertainty of 1.5×10^{-9} L.

12.2.5 Change in phase correction due to wearing of surfaces

The pilot laboratory made measurements at the start and end of the comparison, using the same wrung platens, so any change in length due to changes in phase correction should already be present in the data in section 12.2.1, so no further uncertainty component is proposed.

12.2.6 Additional uncertainties

Although the protocol document requested that all measurement be corrected to the condition of standard atmospheric pressure (101325 Pa), no laboratory performed this correction because for most participants, it is usually very small compared to other corrections. However, as some participants are based at institutes at different altitudes, when comparing results between laboratories at this level of accuracy, the effects of pressure on the gauge block compression should be considered. Strictly, this should be a parameter for each laboratory's uncertainty budget, but for the purposes of this comparison, it will be attributed to an artefact uncertainty (*i.e*. compressibility of the artefact). Data concerning the air pressure at the time of measurement was obtained from the participants and standard equations used to calculate the effect on the gauge block length (see for example H. Darnedde, *Metrologia,* **29**, 349-359 (1992)). The biggest errors for compensation to standard atmosphere that were calculated were 5.3 nm, 15.2 nm, and 27.3 nm, for the 175 mm, 500 mm and 900 mm gauge blocks, respectively. These errors are treated as the full width of a rectangular distribution, leading to standard uncertainties of 1.6 nm, 4.4 nm and 7.9 nm, respectively.

12.2.7 Summary of artefact uncertainties

Combining the additional uncertainties described above, gives the following artefact based standard uncertainties for the four gauge blocks used in CCL-K2:

Serial number	Nominal length (mm)	Artefact-based standard uncertainty (nm)
6071	175	7.3
6071	500	13.0
3701	500	10.3
3701	900	14.1

Table 16 Summary of artefact-based standard uncertainties.

Appendix 2: Comparison with reference values

Table 17 shows the differences Δl of measured lengths with respect to the Key Comparison Reference Values and the expanded (k = 2) uncertainties $U(\Delta l)$ of these differences calculated by

$$U(\Delta l) = 2\sqrt{u^2(x_i) - u^2(\bar{x}_{ref}) + u^2(x_{artefact})} \tag{9}$$

where $u(x_i)$, $u(\bar{x}_{ref})$ and $u(x_{artefact})$ are the standard uncertainties of the laboratory result x_i, the reference value $\bar{x}_{ref}$, and the artefact, $x_{artefact}$.

Laboratory	**175 mm S/N 6071**	**500 mm S/N 6071**	**500 mm S/N 3701**	**900 mm S/N 3701**
IMGC	**-5** ± 57	**-7** ± 69	**-4** ± 67	**+17** ± 84
PTB	-**23** ± 27	**-8** ± 37	**-11** ± 32	**-33** ± 42
NPL	**+16** ± 61	**+39** ± 79	**+43** ± 77	**+41** ± 104
NIST	**-3** ± 33	**-15** ± 50	**-37** ± 48	**+59** ± 120
INMETRO	**+5** ± 41	**+7** ± 45	**+12** ± 42	**+4** ± 70
NRC	**-20** ± 55	**-42** ± 136	**-32** ± 133	**-12** ± 236
NRLM	**+3** ± 39	**+15** ± 82	**+40** ± 79	**+54** ± 136
NIM	**+49** ± 39	**+84** ± 123	**+94** ± 122	**+144** ± 272
CSIRO	**+9** ± 47	**-38** ± 103	**0** ± 101	**-34** ± 174
CSIR	**+35** ± 221	**+57** ± 301	**+52** ± 301	**-6** ± 500
VNIIM	**+167** ± 43	**+29** ± 115	**+50** ± 113	**+149** ± 200

Table 17 Differences of measured lengths with respect to the weighted mean reference values and the expanded uncertainties (k = 2) of these differences (nm). The uncertainties have been rounded up to the nearest nm.

Once again, it should be noted that there are several arguments presented in the main text of this report that indicate that the Key Comparison Reference Values may not be totally reliable estimators of the true values of each measurand, and therefore each laboratory's deviations from these reference values should be interpreted with this in mind.

The calculation of the mutual degrees of equivalence between pairs of laboratories is not recommended for comparisons involving several material standards, since many sets of values would have to be calculated. The mutual degrees of equivalence would be given by $D_{ij} = (x_i - x_j)$ and the expanded uncertainty $U(D_{ij})$ of this difference for two laboratories participating in the same comparison, and by $D_{ij} = (D_i - D_j)$ and its expanded uncertainty $U(D_{ij})$ for two laboratories participating in distinct comparisons, where D_i and D_j are the degrees of equivalence of the two laboratories, in the two different comparisons. However, one can see that Tables 13(a) through 13(d) give the normalized mutual degrees of equivalence, $D_{ij}/U(D_{ij})$ for bilateral comparisons of laboratories' results.

Appendix 3: Transfer of reference values to RMO key comparisons

When trying to link CCL and RMO artefact based key comparisons in dimensional metrology, the application of the concept of transferring the key comparison reference value to a second, independent comparison, turns out to be difficult. Not only does the reference value not have the importance of a realisation of an SI unit but also a rigorous transfer of a numerical reference value would necessitate the introduction of metrologically meaningless corrections and lead to an undue increase in the uncertainty of the regional reference value used to express the degree of equivalence. The RMO and CCL comparisons would then not have equal status, contrary to the expectations of the MRA.

An alternative is to adjudge that the proper link between two comparisons is established by an expert judgement of the results of the participants common to both comparisons, taking into account their degrees of equivalence for all standards of the two comparisons. This follows one of the recommendations of the 2001 CCL-WGDM meeting concerning artefact based key comparisons in dimensional metrology.

Furthermore, the 2002 CCL-WGDM meeting announced that it would formally recommend that artefact based Key Comparisons in Dimensional Metrology would not use a numerical link between the CCL Key Comparison and the corresponding RMO Comparisons. Instead, the link would be based on competences demonstrated by the participants which took part as linking NMIs in the Key and RMO Comparisons. If these linking NMIs were judged to have performed competently in both comparisons then the comparisons were to be regarded as equivalent. The judgement of the competence is the responsibility of the WGDM.