Report on SIM.L-K1 (SIM.4.2) Regional Comparison

Stage One: Calibration of Gauge Blocks by Optical Interferometry

Final Report — 27 September 2006

J. E. Decker[1], J. Altschuler, H. Beladie, I. Malinovsky, E. Prieto, J. Stoup, A. Titov, M. Viliesid, J. R. Pekelsky[1]

[1]Dimensional Metrology Program
Institute for National Measurement Standards (INMS)
National Research Council Canada (NRC)
Ottawa, CANADA K1A 0R6

Abstract

Results of the Stage One portion of the Inter-American System of Metrology (SIM) regional international comparison of gauge block calibration by optical interferometry are presented. In this measurement round-robin, short gauge blocks, 6 made of steel and 6 made of tungsten carbide, in the range of nominal length from 2 mm to 100 mm, were calibrated by 5 national metrology institutes (NMIs) of the SIM region, and one NMI from EUROMET. By employing the technique of optical interferometry, each of the laboratories establishes a direct link to their national primary standard of length through the calibrated laser wavelengths. Results of central length calibration are presented and discussed with regard to vacuum wavelength correction for refractive index of air, phase-change on reflection and wringing effects. Measurement uncertainty evaluation is also discussed.

[1]send correspondence to JED 207 M-36 Montreal Road, NRC, Ottawa, CANADA, K1A 0R6, tel. (613)991-1633, fax. (613)952-1394, e-mail: jennifer.decker@nrc.ca

Contents

1 Introduction

There are several goals associated with an international comparison of measuring artefacts. One goal is to probe the current level of world capability, which then forms the basis of consensus agreement on current state-of-the-art. This international comparison is in the context of absolute central length measurement of gauge blocks.

A benefit to participating laboratories is that each one is able to test its performance. Each participating laboratory has the goal of verifying that their overall measurement system is functioning correctly. This is meaningful because gauge block calibration by optical interferometry involves many sophisticated techniques to form the relationship between the vacuum wavelength of laser-light [1] and the overall mechanical length of a gauge block. These calibrations are a fundamental first step in the chain of traceability to the definition of the metre. Corrections for physical nature of a gauge measuring surface can be elusive placed in the context of wavelengths of light, which are accepted as representing our primary scale with which to establish traceability. International comparison offers the only method to scientifically observe and interrogate the biases that exist in these measurements, even though all the labs are using the same technique, and in some cases even the same instrumentation, and yet also claim direct traceability to the definition of the metre and the ITS-90 temperature scale. And finally, one of the most recent and important goals of international comparison of measurement capability is to support the international Mutual Recognition Arrangement (MRA) [2].

This paper provides a detailed report of the Stage One of a multi-stage international comparison which samples the gauge block calibration service offered to clients by national metrology institutes (NMIs) comprising the SIM region. Stage One comprises the gauge block calibrations by the technique of optical interferometry during the time period of June 1998 to December 1999. The *same* gauge blocks are circulated in Stage Two where SIM NMIs calibrate these gauge blocks using mechanical comparison technique. The same gauge blocks are used successive stages of the comparison, therefore results can shed insight on the link between interferometric and mechanical comparison techniques in addition to all NMIs participating in both stages. A second report outlines the results obtained by mechanical comparison [3]. Several laboratories, namely NIST (USA), INMETRO (Brazil), INTI (Argentina), and CEM (Spain) calibrated the gauge blocks using both techniques thus providing a solid link between the two comparisons.

2 Participants

Out of the six national metrology institutes (NMIs) participating in this comparison, five represent countries of the SIM region and one country represented the EUROMET region. The laboratories, their representative acronyms, and contact information are listed in Table 1. There were many challenges associated with transport and customs issues, however circulation of the gauge blocks to all the labs took about 18 months. The tour circuit for Stage One is outlined in Table 2. INTI, NIST and CEM calibrated the gauge blocks by both interferometry and mechanical comparison techniques, so the time line for these labs is somewhat extended.

The gauge blocks were measured at INMETRO by two completely different gauge block interferometer instruments, also involving different staff members. Data denoted by INMETRO1 represent results from a research-grade instrument for which client calibrations are offered on request. Results denoted by INMETRO2 represent results from the routine interferometric gauge block calibration service offered by INMETRO as listed in the Key Comparison Database (KCDB) Appendix C: Calibration and Measurement Capabilities.

Laboratory	*Contact Information*	*Phone, E-mail*
CENAM	Miguel Viliesid Alonso Metrologia Dimensional Centro Nacional de Metrología (CENAM) Apartado Postal 1-100 Centro 76000 Queretaro Queretaro, Mexico	Tel. +52 42 11 0574 Fax +52 42 11 0577 e-mail: mviliesi@cenam.mx
INMETRO1	C. A. Massone, I. Malinovsky Instituto Nacional de Metrolgia, Normalização e Qualidade Industrial (INMETRO) Av. N. S. das Graas 50 Duque de Caxias, Rio de Janeiro, Brazil	Tel. +55 21 502-1009 Fax. +55 21 293-6559 e-mail: laint@inmetro.gov.br
INMETRO2	Hakima Beladie Instituto Nacional de Metrolgia, Normalização e Qualidade Industrial (INMETRO) Av. N. S. das Graas 50 Duque de Caxias, Rio de Janeiro, Brazil	Tel. +55 21 502-1009 Fax. +55 21 293-6559 e-mail: laint@inmetro.gov.br
INTI	Jeronimo Altschuler Centro de Investigación y Desarrollo en Fisica Instituto Nacional de Tecnologia Industrial (INTI) Parque Tecnolgico Miguelete: Av. General Paz entre Albarellos y Constituyentes CC 157 - (1650) San Martin Buenos Aires, Argentina	Tel. +54 11 4752-5402 Fax +54 11 4713-4140 e-mail: jeroa@inti.gov.ar
NIST	John Stoup National Institute of Standards and Technology (NIST) Room B113, Metrology Building Gaithersburg, MD 20899-0001 USA	Tel. +1 301 975 3476 Fax +1 301 869 0822 e-mail: John.Stoup@nist.gov
CEM	Emilio Prieto Esteban Centro Espanol de Metrologia (CEM) Alfar, 2 - 28760 Tres Cantos Madrid Spain	Tel. +34 91 8074 716 / 700 Fax +34 91 8074 807 e-mail: eprieto@cem.es
NRC (pilot)	Jennifer Decker Institute for National Measurement Standards (INMS) National Research Council Canada (NRC) Ottawa, K1A 0R6, Canada	Tel. +1 613 991 1633 Fax +1 613 952 1394 e-mail: Jennifer.Decker@nrc.ca

Table 1: Participants of SIM.4.2 regional comparison of gauge block calibration, Stage One by optical interferometry.

3 Gauge Block Artefacts

A total of 12 rectangular gauge blocks, ISO 3650 [6] Grade K, were selected. The nominal lengths of the gauge blocks were chosen to provide adequate representation to the range and sampling of short gauge blocks. Moreover, the gauge block nominal lengths echo the short gauge blocks used in the CCL-K1 key comparison. Gauge block materials of steel (CARY, Switzerland) and tungsten carbide (Select, UK) were employed. The nominal lengths of the steel gauge blocks are: 2 mm, 5 mm, 8 mm, 10 mm, 50 mm, 100 mm, and for the tungsten carbide gauge blocks: 2 mm, 5 mm, 8 mm, 20 mm, 50 mm, 100 mm. Gauge blocks were housed in a wooden case. Every attempt was made to hand-carry the gauge blocks whenever possible (July 1998), however, from thereafter were shipped from one laboratory to another because of limited resources and limited travelers between NMIs. Often times shipping took a time duration of one month.

Thermal expansion coefficients for the gauge blocks were not measured, rather the values provided by the manufacturer were requested to be used. These values are: 11.5×10^{-6} /K for steel, and 5.0×10^{-6} /K for tungsten carbide. INMETRO2 used a value of 4.23×10^{-6} /K for tungsten carbide. INMETRO1 measured 4.25×10^{-6} /K for the thermal expansion coefficient of the 100 mm and 50 mm gauge blocks with an uncertainty of 5.0×10^{-8} /K.

Gauge blocks were inspected for damage immediately upon arrival at each laboratory, and a detailed report form outlining the integrity of each gauge block was faxed back to the pilot lab upon receipt of the gauge blocks. Following the first stage of the comparison the gauge blocks were in sufficiently good condition to consider continuing with these gauge blocks for Stage Two of the comparison. Some of the gauge blocks had small scratches, but most of the measuring faces demonstrated good to fair wringing properties following Stage One. The right side of the 50 mm steel gauge block was reported damaged at INMETRO, and was not wringable by INTI or NIST. Indeed, when the pilot lab measured the gauge blocks in October 1999 the right face of the 50 mm steel gauge block was found to be slightly damaged and wringing was compromised. The left face of the 100 mm steel gauge block was difficult to wring following Stage One (it was not wringable by NRC), although INTI, NIST and CEM managed to wring and measure that face.

The gauge blocks appear to be stable in length during the time of the comparison. The pilot lab measured the gauge block at regular intervals in an attempt to monitor the stability of the gauge blocks. During Stage One of this 2-stage comparison, NRC measured the gauge blocks twice. Because of logistical challenges, the pilot lab measured the gauge blocks at a frequency of not more than once per year.

4 Calibration Technique

All participants calibrated the gauge blocks by the technique of optical interferometry, applying the method of exact fractions. Table 3 summarizes details of the equipment used by each laboratory. Most laboratories establish traceability to the definition of the metre through calibration of laser frequency against an iodine-stabilized He-Ne primary standard laser in-house. INTI obtains their traceability through calibration of laser vacuum wavelength by National Physical Laboratory (NPL, UK). Some participating laboratories use lamps as light sources, in particular the cadmium-114 isotope lamp. The electrodeless cadmium-114 lamp correctly operated is considered to be a length standard by the CIPM [1, 12], with absolute accuracy of 7×10^{-8} ($k = 3$ relative uncertainty) in vacuum wavelength, and therefore does not require further calibration to establish traceability to the definition.

The protocol document specified that the gauge blocks are to be calibrated in accordance with the standard

Laboratory	*Dates of Measurement*	*Results Received*
NRC (pilot)	June 1998	—
CENAM	August, September 1998	28 January 1999
INMETRO2	November 1998	11 March 1999
INMETRO1	– February 1999	25 February 1999
INTI	March, April 1999	11 June 1999
NIST	June, July 1999	13 January 2000
NRC (pilot)	October 1999	—
CEM	December 1999 – January 2000	28 April 2000

Table 2: Tour time-line of SIM.4.2 regional comparison of gauge block calibration, Stage One by optical interferometry.

Laboratory	*Instrument & Light Sources*	*Fringe Evaluation*
NRC	NRC Twyman-Green [4, 5] He-Ne lasers 633, 543, 612 nm	Localisation by eye to fiducial in video image
CENAM	Twyman-Green NPL-TESA He-Ne lasers 633, 543 nm	automated DIP of fringe pattern
INMETRO1	Carl Zeiss (modified) He-Ne laser 633 nm	custom DIP of fringe pattern
INMETRO2	Jena-Zeiss 114-Cd lamp 644, 509, 480, 468 nm	visual interpolation
INTI	Twyman-Green NPL-TESA He-Ne lasers 633, 543 nm	automated DIP of fringe pattern
NIST	Fizeau NPL Hilger-Watts (modified) He-Ne laser 633 nm	visual interpolation
CEM	Twyman-Green NPL-TESA He-Ne lasers 633, 543 nm	automated DIP of fringe pattern

Table 3: Summary of instruments and light sources used in gauge block calibrations of the SIM.4.2 Regional Comparison. DIP: Digital Image Processing

ISO 3650 [6], namely that central gauge block length is defined as the height of the centre point of the gauge block measuring face with respect to an auxiliary plane surface. One measuring face of the gauge block is wrung to the auxiliary surface and measured. The gauge block is turned end-over-end and the other measuring face is wrung to the platen and likewise measured. This sequence is repeated so that the result reported by each participant is an average of four separate wringing measurements. In keeping with the ISO 3650 guidelines, certain specific information was requested to be reported by each participant. These data are discussed in turn below.

Following convention of reporting gauge block central length, l is reported as the average of the left and right measuring face wringings, as a deviation d from nominal length L,

$$d = l - L \tag{1}$$

where a plus sign indicates that the gauge block is longer than the nominal length, and a minus sign that it is shorter.

In gauge block interferometry the largest correction for environmental influences is the adjustment of the vacuum wavelength of light for the refractive index of air $\lambda_v = n\lambda_{air}$. All laboratories applied measured values of air temperature, pressure and partial pressure of water vapour to empirical formulae modeling the behaviour of the refractive index of air. Labs varied in the version of the Edlén equation in use. NRC and INTI applied the Birch and Downs 1994 [8] version, CENAM and NIST refer to an update made in 1998 [9]. The other participants did not specify which version was applied. No refractometers were used in this comparison.

5 Results and Discussion

5.1 Central Length Measurement

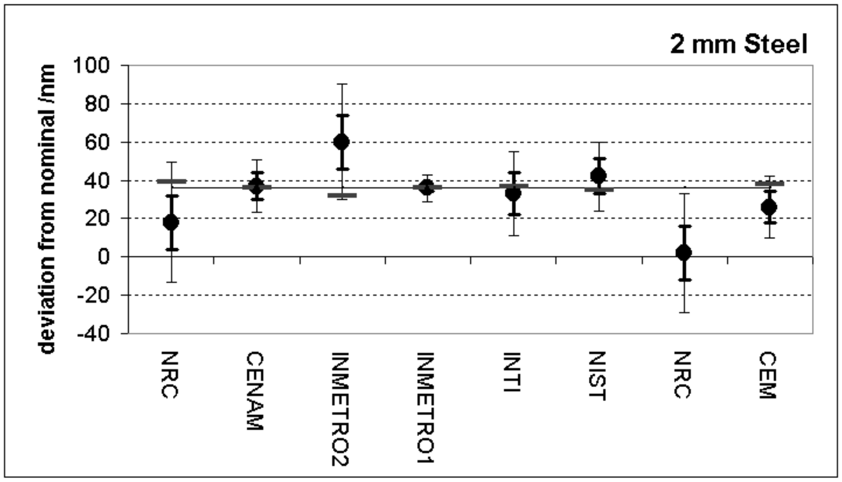
Laboratories submitted detailed reports including: average deviation from nominal central length d, average deviation from nominal length for right and left face wringings, platen materials and phase corrections, standard uncertainty components, combined standard uncertainty and degrees of freedom. Central length measurement values and standard uncertainties reported by each participant are tabulated in Tables 4 through 7, and plotted in Figures 1 and 2 for all gauge blocks of the comparison[2].

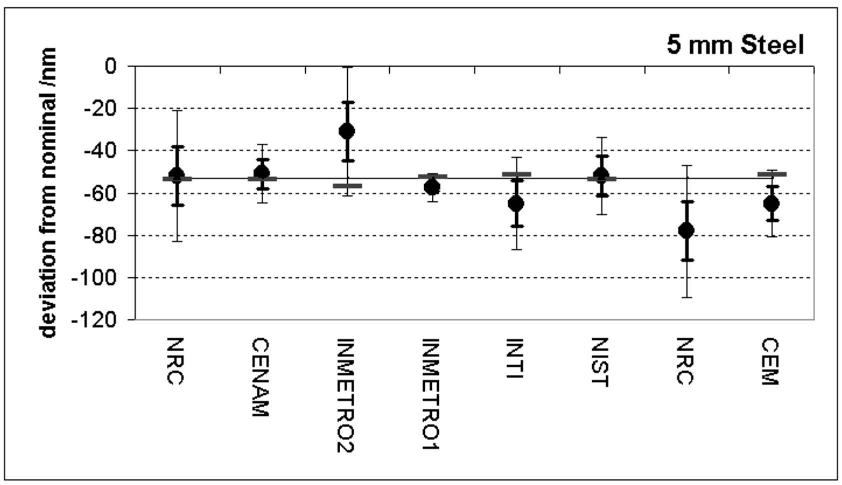
The simple arithmetic mean

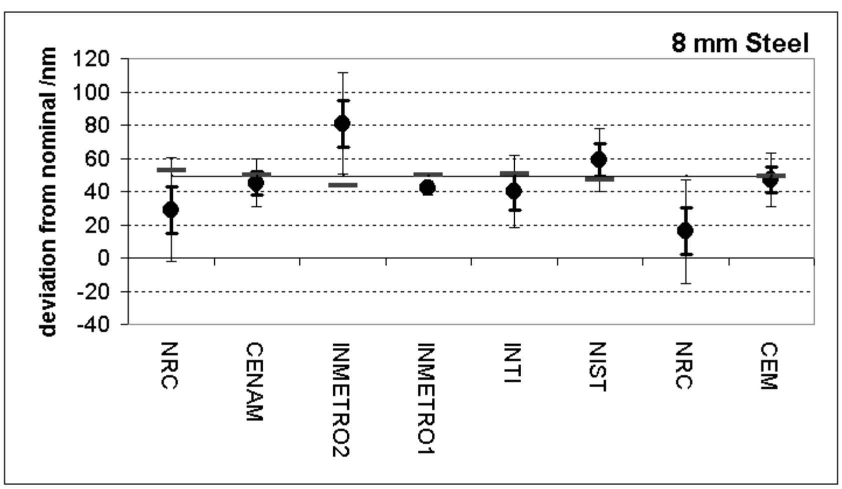
$$\overline{x} = \frac{1}{n}\sum_{x=1}^{n} x_i \tag{2}$$

is included in the data plots, where x_i is the central length measurement reported by each laboratory and n is the number of participants. The NRC pilot measurements intended to probe gauge block stability are included in the plots for information, since there is not enough of this pilot data to warrant a separate plot for Stage One. Only the first measurement of NRC is used in the evaluation of the mean and the KCRV. The exclusive arithmetic mean [10] is also shown in the plots. The exclusive mean is evaluated by taking the mean of all laboratories, leaving out the result of the participant laboratory. This technique allows graphical demonstration of the amount of correlation of each participant with the 'world' simple arithmetic mean.

[2]An oversight in the submission of CENAM was revealed during data tabulation. Measurement results and uncertainties used in the original computations and Draft A Report were confirmed to be correct, and the final Draft B Report remains unchanged from the one accepted by the CCL-WGDM in Sept 2005.


2 mm Steel
deviation from nominal /nm
100
80
60
40
20
0
-20
-40
NRC
CENAM
INMETRO2
INMETRO1
INTI
NIST
NRC
CEM


5 mm Steel
deviation from nominal /nm
0
-20
-40
-60
-80
-100
-120
NRC
CENAM
INMETRO2
INMETRO1
INTI
NIST
NRC
CEM


8 mm Steel
deviation from nominal /nm
120
100
80
60
40
20
0
-20
-40
NRC
CENAM
INMETRO2
INMETRO1
INTI
NIST
NRC
CEM

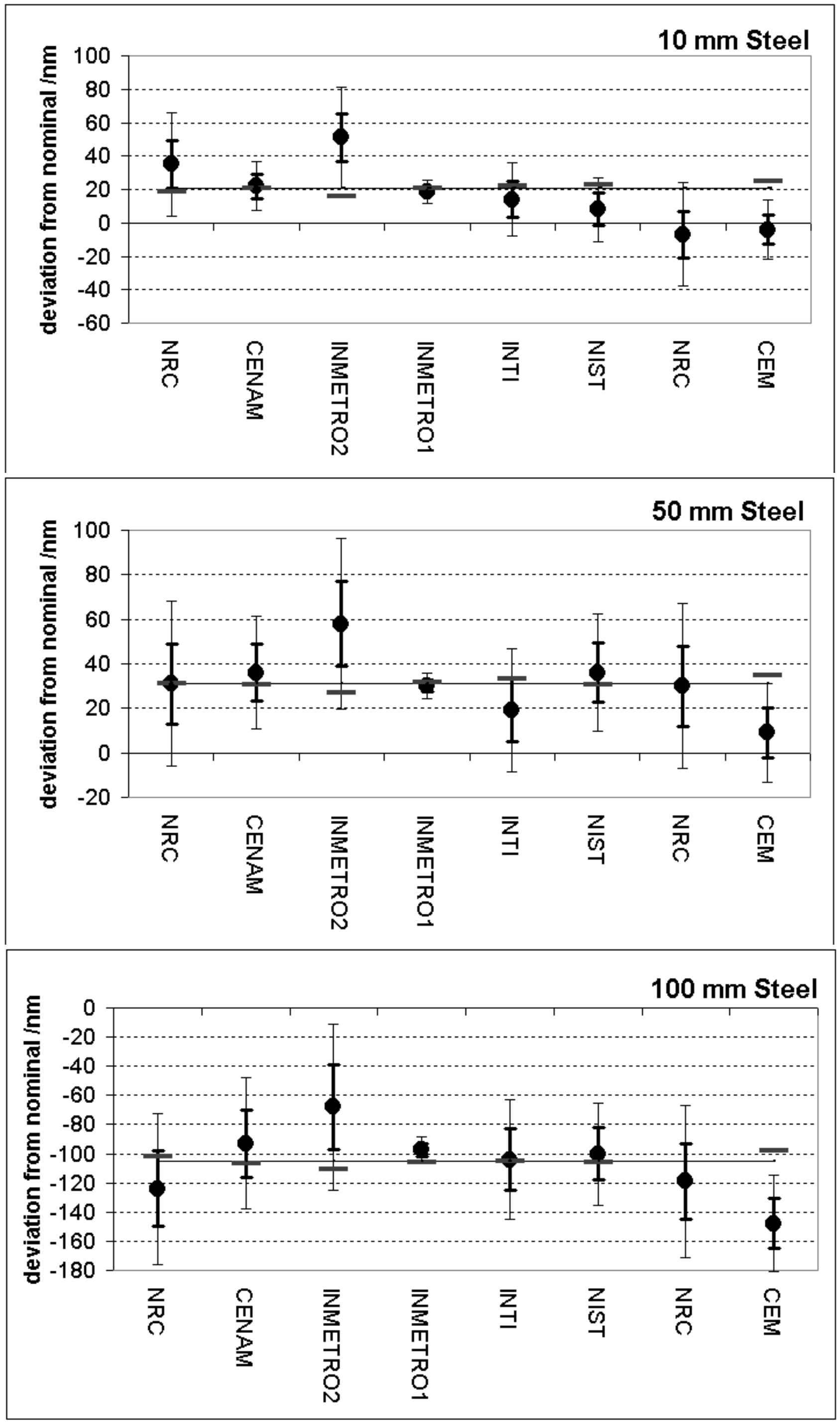


Figure 1: Plot of central length expressed as deviation from nominal length reported by each participant for steel gauge blocks. Thick error bars represent the standard uncertainty, while longer thin error bars represent $k_{95}u(x_i)$ where $k_{95} = t_p(\nu_i)$ from the Student's t-distribution for standard uncertainties $u(x_i)$ and degrees of freedom ν_i submitted by the participants. The solid line represents the simple arithmetic mean of the reported central lengths. The dash for each participant represents the exclusive simple arithmetic mean (see text).

Nominal Length /mm	Deviation from Nominal Length for Steel Gauge Blocks /nm						
	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
2	18	37	60	36	33	42	26
5	-52	-51	-31	-58	-65	-52	-65
8	29	45	81	42	40	59	47
10	35	22	51	19	14	8	-4
50	31	36	58	30	19	36	9
100	-124	-93	-68	-98	-104	-100	-148

Table 4: Central length expressed as deviation from nominal length reported by each participant for steel gauge blocks.

Nominal Length /mm	Standard Uncertainties for Steel Gauge Blocks /nm						
	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
2	14	7	14	2	11	9	8
5	14	7	14	2	11	9	8
8	14	7	14	2	11	10	8
10	14	7	14	2	11	10	9
50	18	13	19	3	14	13	11
100	26	23	29	4	21	18	17

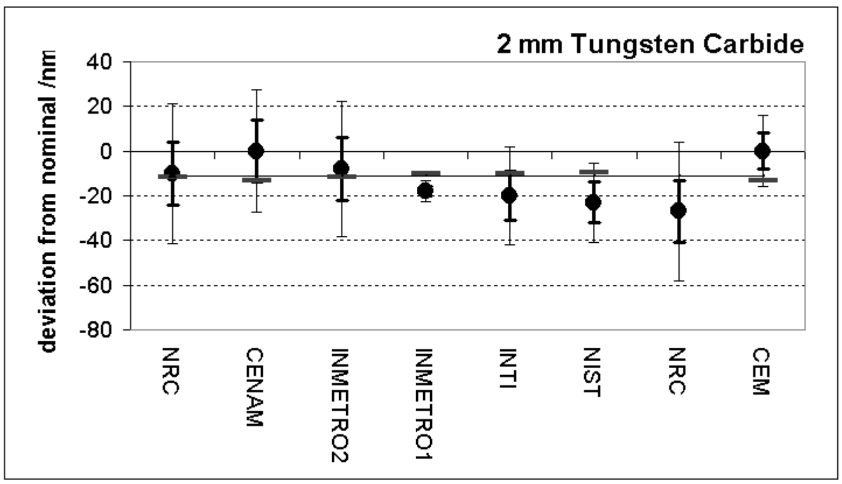
Table 5: Combined standard uncertainty attributed to steel gauge block central length measurement as reported by each participant.

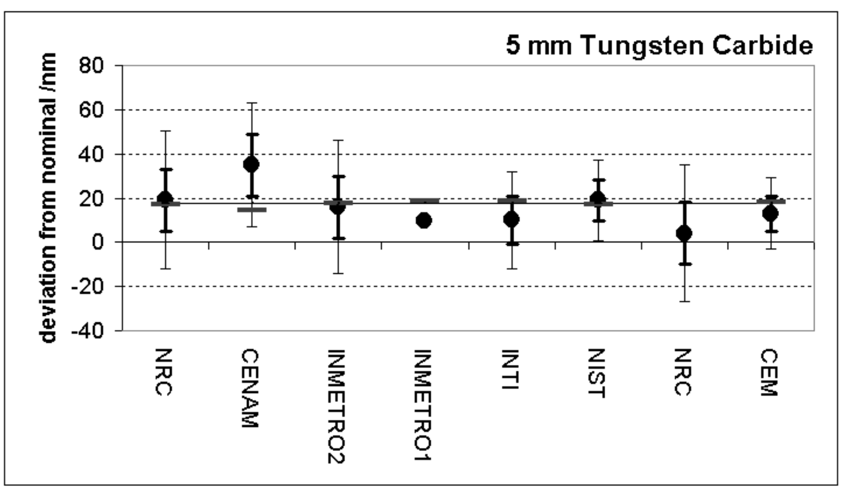
Nominal Length /mm	Deviation from Nominal Length for Tungsten Carbide Gauge Blocks /nm						
	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
2	-10	0	-8	-18	-20	-23	0
5	19	35	16	10	10	19	13
8	45	54	34	30	28	31	39
20	10	21	1	-5	-2	18	17
50	-25	-13	-24	-40	-36	-32	-43
100	-58	-27	-28	-36	-57	-46	-63

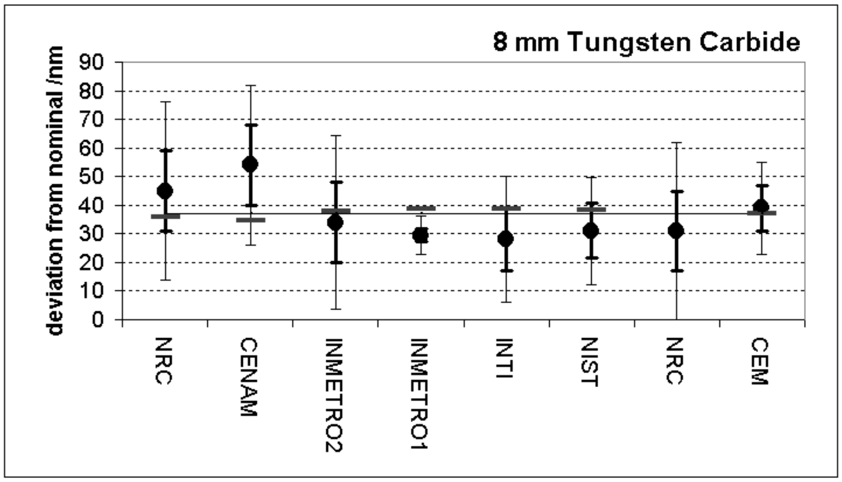
Table 6: Central length expressed as deviation from nominal length reported by each participant for tungsten carbide gauge blocks.

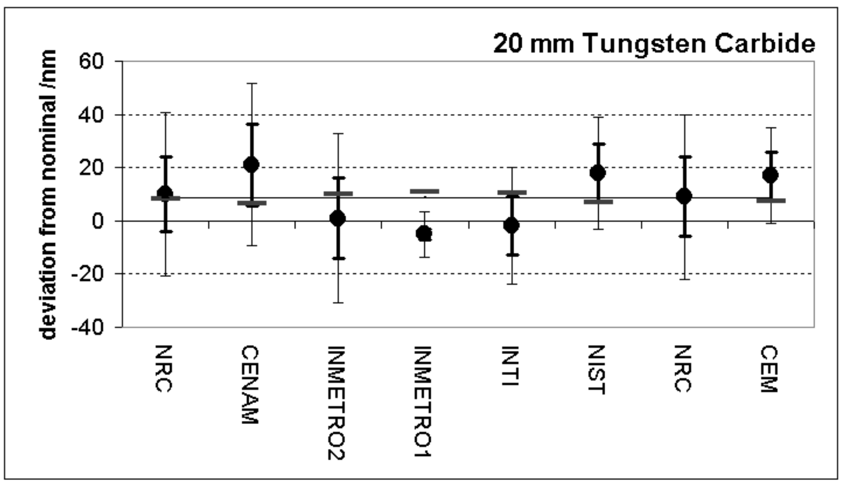
Nominal Length /mm	Standard Uncertainties for Tungsten Carbide Gauge Blocks /nm						
	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
2	14	14	14	2	11	9	8
5	14	14	14	1	11	9	8
8	14	14	14	1	11	10	8
20	14	15	15	2	11	11	9
50	17	22	18	3	12	13	10
100	24	36	26	4	16	18	13

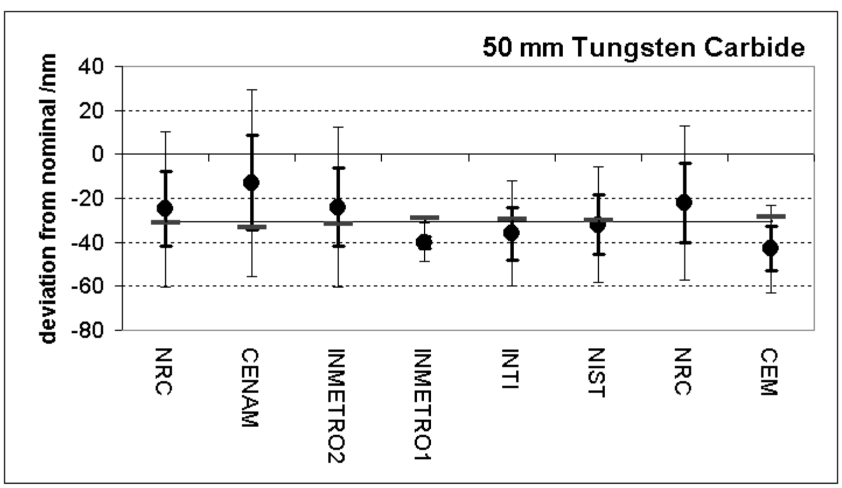
Table 7: Combined standard uncertainty attributed to tungsten carbide gauge block central length measurement reported by each participant.






8 mm Tungsten Carbide
deviation from nominal /nm
90
80
70
60
50
40
30
20
10
0
NRC
CENAM
INMETRO2
INMETRO1
INTI
NIST
NRC
CEM


20 mm Tungsten Carbide
deviation from nominal /nm
60
40
20
0
-20
-40
NRC
CENAM
INMETRO2
INMETRO1
INTI
NIST
NRC
CEM


50 mm Tungsten Carbide
deviation from nominal /nm
40
20
0
-20
-40
-60
-80
NRC
CENAM
INMETRO2
INMETRO1
INTI
NIST
NRC
CEM

		Difference Between Left and Right Measuring Face Wringing /nm						
		NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
Steel	maximum	17	9	20	12	2	15	17
	st dev	9	6	11	7	1	9	9
Tungsten Carbide	maximum	16	15	31	13	12	16	17
	st dev	10	8	11	6	7	8	6

Table 8: Maximum difference between central length measurements of left and right wringings listed with the standard deviation of this difference for all gauge blocks in each material sample (nm units).

5.2 Difference Between Left and Right Measurement Face Wringing

The protocol requested that participants report the average d for right and left measuring face wringings. Figures 3 and 4 plot these results for steel and tungsten carbide gauge blocks respectively. Differences in length measurements between left and right side wringings can indicate a geometry feature of the gauge block that results in different wringing qualities between left and right. The quality of the gauge block, the platen and the technical experience of the metrologist all influence the the closeness of left and right wring length measurements. Participants of this comparison showed similar results in this category.

5.3 Phase Correction

According to the ISO 3650 definition of gauge block length, the central length measurement must include the appropriate corrections for difference in material or surface texture between the platen and the gauge block measuring face [6, 7]. INMETRO1 used their technique of reproducible wringing [11], whereas all other participants used stack techniques for the evaluation of their correction for phase-change on reflection. INMETRO2 did not perform pack experiments on the steel gauge blocks as they did on the tungsten carbide gauge blocks, but rather applied a phase correction for steel based on previous characterization experiments. The comparison protocol included specific instructions for reporting these phase change correction values. Submitted values are listed in Table 9.

5.4 Measurement Uncertainty

To expedite analysis of comparison comparison results, labs were requested to provide a measurement uncertainty budget in the model of that described in [13], including the standard uncertainty components attributed to the largest influence parameters of their calibration. An example table was provided in the protocol document.

Each laboratory submitted a summary evaluation of their measurement uncertainty. The following influence parameters were identified in the submitted uncertainty evaluations:

- λ_i: vacuum wavelength of the light sources,
- F_i: measurement of interference fringe fraction,

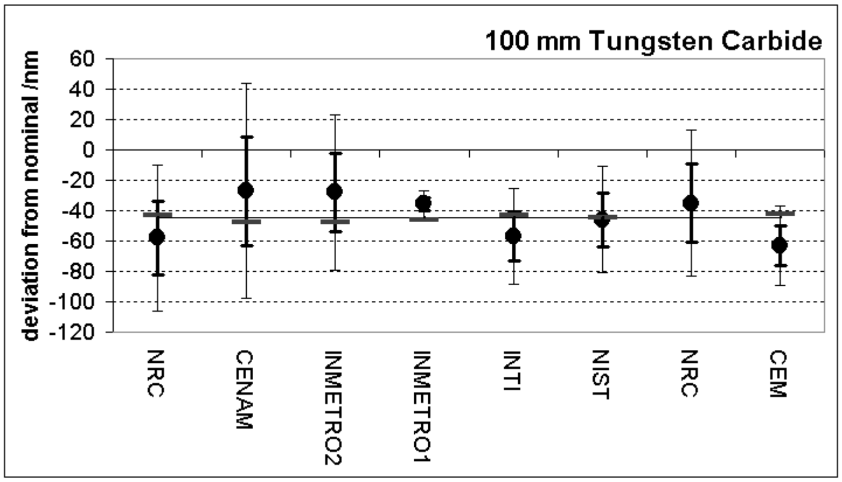


Figure 2: Plot of central length expressed as deviation from nominal length reported by each participant for tungsten carbide gauge blocks. Thick error bars represent the standard uncertainty, while longer thin error bars represent $k_{95}u(x_i)$ where $k_{95} = t_p(\nu_i)$ from the Student's t-distribution for standard uncertainties $u(x_i)$ and degrees of freedom ν_i submitted by the participants. The solid line represents the simple arithmetic mean of the reported central lengths. The dash for each participant represents the exclusive simple arithmetic mean (see Section 5.1 text).

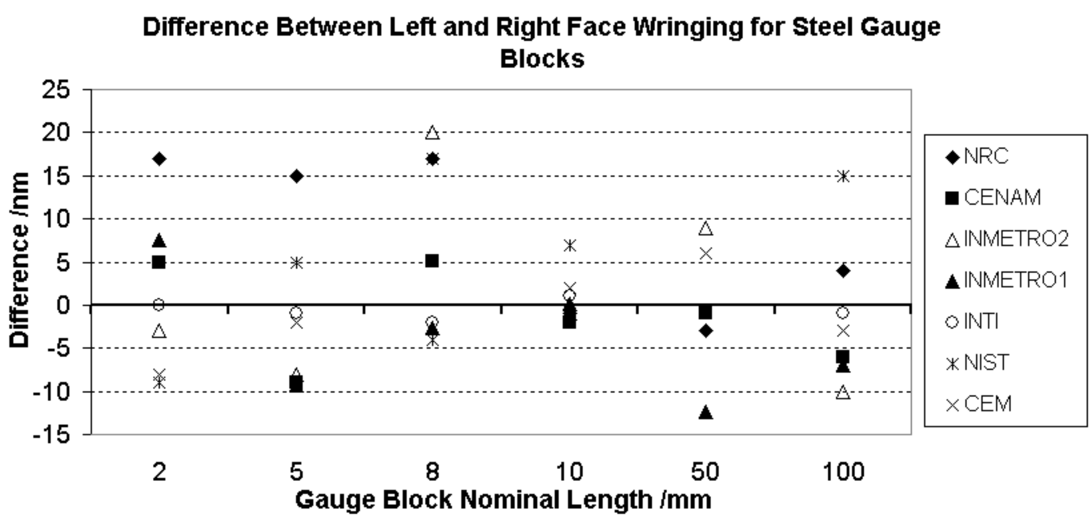


Figure 3: Plot of differences between left and right wringing compared to average central length for steel gauge blocks.

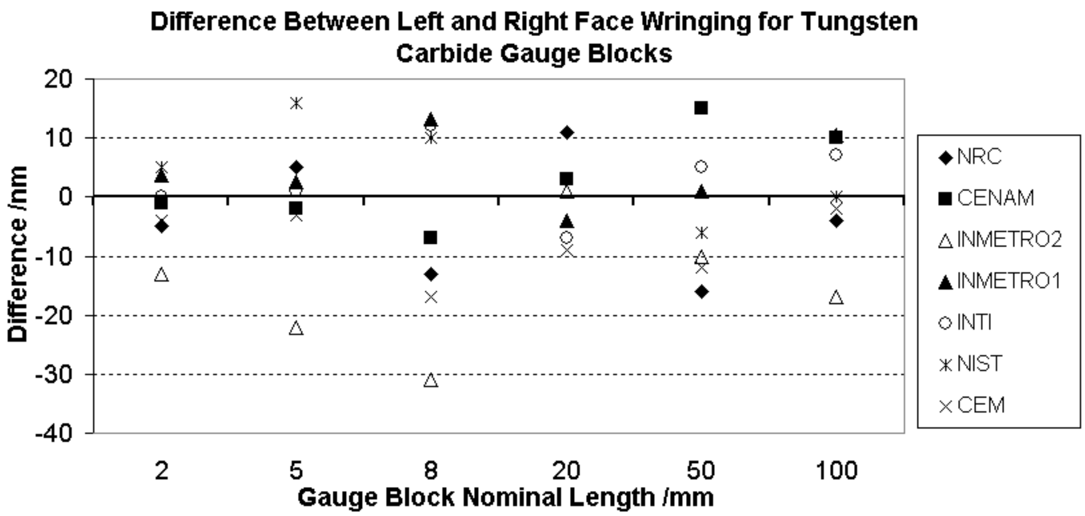


Figure 4: Plot of differences between left and right wringing compared to average central length for each participant, for tungsten carbide gauge blocks.

Laboratory	*Steel Gauge Blocks*		*Tungsten Carbide Gauge Blocks*	
	Platen Material	*Phase Correction /nm*	*Platen Material*	*Phase Correction /nm*
NRC	fused silica	+51	fused silica	+43
CENAM	steel (TESA, UK)	−23	tungsten carbide (TESA, UK)	−20
INMETRO1	steel (Cary, CH)	'slave block'	steel	'slave block'
INMETRO2	quartz	+45	quartz	+19
INTI	steel (TESA, UK)	−46	tungsten carbide (TESA, UK)	−39
NIST	steel	+11	steel	−11.7
CEM	steel (TESA, UK)	−17	tungsten carbide (TESA, UK)	+13

Table 9: Summary of platen materials and phase corrections for steel and tungsten carbide gauge blocks.

- n: refractive index of air (combined standard uncertainty includes components of air temperature, air pressure and relative humidity measurements),
- Δt_g: gauge block temperature measurement,
- α: linear coefficient of thermal expansion,
- δl_Ω: obliquity correction – alignment of the entrance aperture,
- Δl_s: obliquity correction – size of the source aperture,
- δl_A: wavefront aberrations,
- δl_G: departure from perfect prismatic geometry of the gauge block,
- δl_w: wringing,
- Δl_ϕ: phase correction (combined standard uncertainty).

Combined standard uncertainty values reported for each gauge block in the comparison are listed in Tables 3 and 5 for the steel and tungsten carbide gauge blocks respectively. Table 10 provides a general summary of the range of expanded uncertainty and degrees of freedom for the steel gauge blocks, for nominal gauge block lengths 2 mm to 100 mm. Individual components of standard uncertainty reported by each participant are listed in Table 11.

Two of the largest influences on gauge block calibration by optical interferometry are air pressure and temperature which in turn, directly influence refractive index of air n. Uncertainty components for these influences are nested in the total uncertainty for refractive index. Temperature also affects uncertainties related to the thermal expansion of the gauge block through Δt_g.

The histogram of the pooled comparison data shown in Figure 5 demonstrate that comparison results follow a distribution similar to the normal distribution. More importantly, examination of the data plots and the histograms demonstrate that there are no obvious outliers in the comparison data. Outlier data points could affect the evaluation of the reference value in a detrimental way by falsely pulling the mean in the direction of the outlier data point. Therefore all data submitted by the participants can be used in the evaluation of the reference values.

6 Conclusions

Results of SIM.4.2 regional comparison of gauge block calibration by optical interferometry are reported. Data are presented in the form of tables and plots of deviation from nominal length reported by each participating laboratory for each gauge block of the comparison. Comparison data for left and right wringing differences, uncertainty evaluations, and equipment styles are also reported. Gauge block nominal lengths and materials were selected to probe the range of nominal lengths between 2 mm and 100 mm.

The simple arithmetic mean of the central length measurement is evaluated for each gauge block in the comparison, and is recommended as the comparison reference value (KCRV). Tables in the Appendix list the difference between the individual result of each participant with respect to the KCRV, along with the expanded uncertainty of this difference. Tables of bilateral equivalence are also included in the Appendix.

Laboratory	*Maximum Temperature Variation During Measurements* /°C	*Range of Expanded Uncertainty (Steel)* /nm	*Range of Degrees of Freedom (Steel)*
NRC	±0.03	28 – 52	10 – 73
CENAM	±0.25	14 – 46	72 – 251
INMETRO1	±0.15	3 – 9	2 – 21
INMETRO2	±0.25	28 – 58	13 – 256
INTI	±0.1	22 – 42	75 – 1086
NIST	+0.2	18 – 36	—
CEM	±0.05	16 – 34	71 –261

Table 10: Summary of details regarding temperature range during measurements, reported expanded uncertainties and degrees of freedom for range 'boundary' values of 2 mm and 100 mm nominal (steel) gauge block lengths.

	Components of Standard Uncertainty /nm						
	NRC	CENAM	INMETRO1	INMETRO2	INTI	NIST	CEM
λ_i	*0.5*	*3*	*0.2*	*1.5*	*3*	*0.3*	*1.1*
F_i	2	3	0.1	6.6	3	4.5	4.2
n	*20*	*2.5*	*3.4*	*15.3*	*8.9*	*3*	*7.5*
Δt_g	*7.2*	*16.7*	*1.7*	*12.9*	*13.8*	*8.5*	*10*
α	*3*	*15*	*0.75*	*15*	*5.7*	*0.8*	*3.5*
δl_Ω	*0.8*	*0.6*	*0.3*	—	*0.6*	*0.1*	*0.6*
Δl_s	*0.2*	*0.4*	—	*0.5*	*0.3*	—	*0.2*
δl_A	3	3.5	0.3	3	3.4	3	3
δl_G	3	1.4	0.1	2	1.4	—	2.5
δl_w	8	3	1.5	6	7	4	3
Δl_ϕ	10	4	0.5	10	6	5.8	5
u_c	**25**	**24**	**4**	**29**	**21**	**13**	**15**

Table 11: Summary of components of standard uncertainty for gauge block calibration by interferometry reported by participants of the SIM.4.2 Regional Comparison. Length dependent terms are in italics and are based on 100 mm nominal gauge block length.

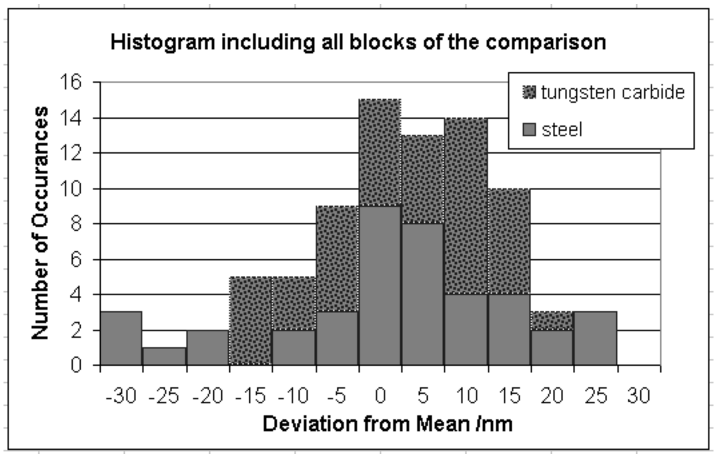


Figure 5: Histogram combining all gauge blocks of the comparison. Number of occurrences of difference value from the simple arithmetic mean.

Participating laboratories demonstrate general agreement in measurement of central gauge block length within the average scatter of data around ±25 nm. Comparison data is tighter for the tungsten carbide gauge block samples, likely owing to the more reproducible wringing quality of these blocks. The surface characteristics of the tungsten carbide gauge blocks are better for wringing than steel, and the material is more durable in a comparison exercise.

Dispersion observed in the values of phase correction reported by participants using the same platens from the same manufacturers invite further investigation into the variation of phase correction values even while employing the same equipment.

Stage One of this comparison took a total of 18 months to complete. The gauge blocks returned in reasonable condition, deemed sufficiently good to continue to Stage Two of the SIM.4.2 regional gauge block comparison [3]. In Stage Two, participants calibrate the same gauge blocks by mechanical comparison methods.

Three of the seven gauge block interferometer instruments were identical instruments purchased from the same manufacturer. Agreement is demonstrated between laboratories with automated fringe evaluation, particularly for gauge block measurements of shorter nominal length.

In future comparisons, it would be advantageous to report detailed measurement uncertainty of air temperature and pressure measurement rather than overall refractive index. The influence of air pressure and temperature on vacuum wavelength corrections could have correlation with comparison results.

This comparison provides a link to the CCL-K1 Key Comparison of short gauge block calibration by interferometry through NIST, NRC and CENAM. This Comparison also provides a link to the EUROMET Comparison of short gauge block calibration through CEM, and to the SIM.4.2 Stage Two Short Gauge Block Comparison by Mechanical Comparison through NIST, INTI, CEM, and INMETRO.

7 Acknowledgements

The SIM.4.2 Regional Gauge Block Comparison participants gratefully acknowledge the financial support of the SIM regional committee. NRC acknowledges the technical support of M. Dagenais for gauge block calibration work, and L. Munro for maintenance of the NRC Gauge Block Interferometer Laboratory. CENAM would like to acknowledge Juan Carlos Zarraga and Carlos Colin for performing the careful measurements for this comparison. INTI would like to acknowledge Jorge Alvarez and Sergio Ilieff for having taken part in the measurements. CEM would like to acknowledge Joaquín Rodriguez for performing both interferometric and mechanical comparison measurements for this comparison. The author also thanks NRC-INMS colleagues Alan Steele, Rob Douglas and Barry Wood for fruitful discussions.

References

[1] Quinn, T.J., 2003, "Practical realization of the definition of the metre, including recommended radiations of other optical frequency standards (2001)," *Metrologia*, **40**, 103–133.

[2] http://www.bipm.fr

[3] Decker, J.E., Alschuler J., Castillo Candanedo, J., De la cruz, E., Prieto Esteban, E., Morales, R., Valente de Oliveira, J.C., Stone, J., Stoup, J., Pekelsky, J.R., 2003, "SIM.4.2 Regional Comparison Stage Two: Calibration of gauge blocks by mechanical comparison," *Metrologia*, **40** No. 1A, 04003.

[4] Decker, J.E. and Pekelsky, J.R., "Gauge Block Calibration by Optical Interferometry at the National Research Council Canada," presented at the Measurement Science Conference, Pasadena, California, 23–24 January 1997, NRC Document No. 40002.

[5] Decker J.E., Bustraan, K., de Bonth, S., Pekelsky, J.R., "Updates to the NRC gauge block interferometer," NRC Document No. 42753, August 2000, 9 pages.

[6] ISO 3650:1998(E), 1998, 2^{nd} Edition 1998-12-15, (International Organization for Standardization (ISO) Central Secretariat, Switzerland, Internet iso@iso.ch).

[7] Leach, R.K., Hart, A., 2000, "EUROMET Project 413: Interlaboratory comparison of measurements of the phase correction in the field of gauge block interferometry," *Metrologia*, **37**, 261–267.

[8] Birch, K.P. and Downs, M.J., "Correction to the Updated Edlén Equation for the Refractive Index of Air," *Metrologia*, **31**, 315–316 (1994).

[9] Bönsch, G., Potulski, E., 1998, "Measurement of the refractive index of air and comparison with modified Edlén's formulae" *Metrologia*, **35**, pp. 133–139.

[10] Steele, A.G., Wood, B.,M., Douglas, R.,J., 2001, "Exclusive statistics: simple treatment of the unaviodable correlations from key comparison reference values," *Metrologia*, **38**, 483–488.

[11] Titov, I., Malinovsky, A., Balaidi, H., Franca, R.S., Massone, C.A., 1998, "Sub-nanometer Precision in Gauge Block Measurements," CPEM 98, July 1998, Washington DC, USA.

[12] *Comité Consultatif pour la Définition du Mètre*, 3rd Session, 1962, pp. 18-19 and *Procès-Verbaux CIPM*, 52nd Session, 1963, pp. 26-27.

[13] Decker, J.E. and Pekelsky, J.R., 1997, "Uncertainty Evaluation for the Measurement of Gauge Blocks by Optical Interferometry," *Metrologia*, **34**, 479–493. [NRC Document No. 41374]

[14] Cox, M.G., 2002, "The evaluation of key comparison data," *Metrologia*, **39**, 589–595.

[15] Beissner, K., 2002, "On a measure of consistency in comparison measurements," *Metrologia*, **39**, 59–63.

[16] Beissner, K., 2003, "On a measure of consistency in comparison measurements: II Using effective degrees of freedom," *Metrologia*, **40**, 31–35.

[17] Kacker, R., Datla, R., Parr, A., 2002, "Combined result and associated uncertainty from interlaboratory evaluations based on the ISO Guide," *Metrologia*, **39**, 279–293.

[18] Bevington, P.,R., Data Reduction and Error Analysis for the Physical Sciences, (McGraw-Hill, New York 1969) p. 84.

[19] Birge, R.,T., 1932, "The Calculation of Errors by the Method of Least Squares," *Phys. Rev.*, **40**, 207–227.

[20] Taylor, B.,N., Parker, W.,H., Langenberg, D.,N., 1969, "Determination of e/h, QED, and the Fundamental Constants," *Rev. Mod. Phys.*, **41**, 375–496.

[21] Thalmann, R., 2002, "CCL key comparison: calibration of gauge blocks by interferometry," *Metrologia*, **39**, 165–177.

[22] *Guide to the Expression of Uncertainty in Measurement,* 2nd ed., Geneva, International Organization for Standardization, 1995.

A Evaluation of the Comparison Reference Value

The Key Comparison Reference Value (KCRV) x_{ref} and tables of equivalence are evaluated following some recommendations found in the metrology literature [14, 15, 16, 17]. The following discussion first considers the most reasonable mean value to apply for the SIM.4.2 Stage One KCRV, followed by tables of equivalence and bilateral equivalence.

A.1 Evaluations and Calculations

The inverse-variance weighted mean is evaluated by the following equation:

$$y = \frac{\sum_{i=1}^{n} x_i \, u^{-2}(x_i)}{\sum_{i=1}^{n} u^{-2}(x_i)} \tag{3}$$

where x_i are the measurement results and $u(x_i)$ the standard uncertainties submitted by the comparison participants, and n represents the number of participants contributing to the evaluation of the mean. The standard uncertainty $u(y)$ associated with y is

$$u(y) = \frac{1}{\sqrt{\sum_{i=1}^{n} u^{-2}(x_i)}}. \tag{4}$$

The uncertainty in the arithmetic mean of equation (2) can be expressed as

$$u(\overline{x}) = \frac{1}{n}\sqrt{\sum_{i=1}^{n} u^2(x_i)}. \tag{5}$$

Expression of equivalence d_i typically takes the form of the difference between the participant's measured value x_i and x_{ref}.

$$d_i = x_i - x_{\mathrm{ref}} \tag{6}$$

The selection of the KCRV from amongst the inverse-variance weighted mean of equation (3), the simple arithmetic mean of equation (2), or the median usually depends on the overall consistency of the data sets. Taking an analogy to curve fitting, one has more confidence in a fit if the data does not contain outliers, or data that somehow creates a dominating influence. For the purpose of providing scientific evidence to support the CIPM-MRA, the inverse-variance weighted mean has the advantage that both participant measurement values *and* their evaluations of standard uncertainties are probed in the tests for consistency.

Statistical consistency of a comparison can be checked by evaluating the observed chi-squared value [14, 18]

$$\chi^2_{\mathrm{obs}} = \sum_{i=1}^{n} \frac{(x_i - y)^2}{u^2(x_i)}. \tag{7}$$

The consistency check fails if

$$\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\} < 0.05. \tag{8}$$

For the seven participants in this part of the comparison, the degrees of freedom $\nu = 7 - 1 = 6$. Values of the calculated probabilities are listed in the Tables below. Similarly, if the variance weighted mean is a good representation of the data, then the value of the reduced chi-squared $\chi^2_\nu = \chi^2_{\mathrm{obs}}/\nu$ should be approximately unity $\chi^2_\nu = 1$. With regard to the relative significance of these tests, it is important to consider that chi-squared tests are valid only if all participant distributions are Gaussian with mean value equal to the participant's stated value x_i, and standard deviation equal to $u(x_i)$.

Another metric for evaluation of statistical consistency is the Birge Ratio [17, 19, 20], defined as

$$R_B = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n} w_i\,(x_i - y)^2} \tag{9}$$

where the weights $w_i = 1/u^2(x_i)$ for $i = 1, \ldots, n$ are evaluated from the self-declared standard uncertainties. Consistency means that the results x_i and the standard uncertainties $u(x_i)$ fit the Birge Ratio model, which in turn means that values of R_B that are close to 1 or less suggest that the results of the comparison are consistent. Values of R_B that are much greater than 1 suggest that results x_i are inconsistent. Since the Birge Ratio calculation includes $u(x_i)$ as known parameters representing standard deviations of lab results x_i, the Birge Ratio test requires that each of the uncertainties be reliable. When this assumption is not well justified, the conclusion of the Birge ratio test should not be taken too seriously. This warning, and the one stated above for chi-squared, are particularly relevant considering the very low degrees of freedom stated for INMETRO1 for short gauge block nominal lengths.

Another simple method to probe the consistency of a data set is to confirm that

$$|d_i| < k_{95}\; u(x_i) \tag{10}$$

for all participants [15]. Values of d_i listed in the Tables apply the simple arithmetic mean as the reference value. In this comparison, the stated degrees of freedom from each participant do not rigorously allow for a coverage factor of $k = 2$ at a level of confidence of about 95 % for all laboratories. Therefore in the Tables below, individual k_{95} values are evaluated from the Student's t-distribution taking into consideration the participant's submitted ν_i.

The standard uncertainty $u(d_i)$ in the stated equivalence d_i in the case where x_{ref} is evaluated by the arithmetic mean equation (2), and $u(x_{\mathrm{ref}})$ by equation (5) is expressed by [15] (see also [21]):

$$u(d_i) = \sqrt{u^2(x_i) + u^2(x_{\mathrm{ref}}) - \frac{2}{n}u^2(x_i)}. \tag{11}$$

This expression is used because it considers the correlation of each lab with the mean value. Even though it can be shown through exclusive statistics that the amount of correlation between participant labs and the simple arithmetic mean is small (see data plots in Figures 1 and 2 and discussion below) the more general approach is taken here.

The normalized deviation and its consistency limit are then calculated following

$$E_i = \frac{d_i}{k\; u(d_i)} \quad \text{and} \quad |E_i| < 1, \tag{12}$$

where $k = 2$ represents a confidence level of approximately 95 % that the measured value is within $\pm U$ of the true value (for a normal distribution). A more thorough approach [16] evaluates

$$E_{95,i} = \frac{d_i}{k_{95}(d_i)\; u(d_i)} \quad \text{and} \quad |E_{95,i}| < 1 \tag{13}$$

where the self-declared degrees of freedom are used with the Welch-Satterthwaite equation to determine effective degrees of freedom for x_{ref}. Coverage factors $k_{95}(d_i)$ are evaluated from $\nu_{\mathrm{eff}}(d_i)$ and the Student's t-distribution (see [16, 22] for detail). Both versions of normalized deviation are tabulated with the other indicators of statistical consistency mentioned above.

100 mm Tungsten Carbide				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	-13	48	0.30	0.30
CENAM	18	71	0.29	0.29
INMETRO2	17	51	0.37	0.36
CEM	-18	26	0.67	0.66
INTI	-12	32	0.39	0.38
NIST	-1	35	0.03	0.03
INMETRO1	9	9	0.53	0.52

Median	**$-$46.0 nm**
Simple arithmetic mean:	**$-$44.9 nm**
standard uncertainty	8.2 nm
Variance weighted mean:	**$-$39.5 nm**
standard uncertainty	3.6 nm
Variance weighted mean sans INMETRO1:	**$-$53.3 nm**
standard uncertainty	7.7 nm
Observed chi-squared	**6.4**
Observed chi-squared sans INMETRO1	2.3
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$	0.38
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ sans INMETRO1	0.89
Reduced chi-squared	**1.07**
sans INMETRO1	0.38
Birge Ratio	**1.03**
sans INMETRO1	0.62

50 mm Tungsten Carbide				
	d_i	$k_{95}u(x_i)$	$\|E_{95,i}\| = d_i/k_{95}(d_i)u(d_i)$	$\|E_i\| = d_i/2u(d_i)$
NRC	5	35	0.17	0.18
CENAM	17	43	0.46	0.46
INMETRO2	6	36	0.20	0.20
CEM	-13	20	0.63	0.62
INTI	-6	24	0.24	0.24
NIST	-2	26	0.06	0.06
INMETRO1	-10	9	0.81	0.80

Median	**$-$32.0 nm**
Simple arithmetic mean:	**$-$30.4 nm**
standard uncertainty	5.6 nm
Variance weighted mean:	**$-$38.8 nm**
standard uncertainty	2.4 nm
Variance weighted mean sans INMETRO1:	**$-$33.5 nm**
standard uncertainty	5.7 nm
Observed chi-squared	**3.4**
Observed chi-squared sans INMETRO1	2.4
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$	0.75
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ sans INMETRO1	0.88
Reduced chi-squared	**0.57**
sans INMETRO1	0.40
Birge Ratio	**0.76**
sans INMETRO1	0.63

20 mm Tungsten Carbide				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	1	31	0.05	0.06
CENAM	12	30	0.46	0.45
INMETRO2	−8	32	0.27	0.28
CEM	8	18	0.48	0.48
INTI	−11	22	0.52	0.51
NIST	9	21	0.48	0.47
INMETRO1	−14	9	1.43	1.42

Median	**10.0 nm**
Simple arithmetic mean:	**8.6 nm**
standard uncertainty	4.5 nm
Variance weighted mean:	**−2.6 nm**
standard uncertainty	1.8 nm
Variance weighted mean sans INMETRO1:	**11.5 nm**
standard uncertainty	4.8 nm
Observed chi-squared	**13.1**
Observed chi-squared sans INMETRO1	3.1
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$	0.04
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ sans INMETRO1	0.79
Reduced chi-squared	**2.18**
sans INMETRO1	0.52
Birge Ratio	**1.48**
sans INMETRO1	0.72

8 mm Tungsten Carbide				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	8	31	0.28	0.31
CENAM	17	28	0.67	0.67
INMETRO2	−3	30	0.12	0.13
CEM	2	16	0.11	0.11
INTI	−9	22	0.46	0.45
NIST	−6	19	0.35	0.34
INMETRO1	−8	7	0.85	0.84

Median	**34.0 nm**
Simple arithmetic mean:	**37.2 nm**
standard uncertainty	4.2 nm
Variance weighted mean:	**30.9 nm**
standard uncertainty	1.9 nm
Variance weighted mean sans INMETRO1:	**37.1 nm**
standard uncertainty	4.5 nm
Observed chi-squared	**5.3**
Observed chi-squared sans INMETRO1	3.0
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$	0.50
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ sans INMETRO1	0.81
Reduced chi-squared	**0.89**
sans INMETRO1	0.49
Birge Ratio	**0.94**
sans INMETRO1	0.70

5 mm Tungsten Carbide				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	2	31	0.06	0.06
CENAM	18	28	0.70	0.70
INMETRO2	-1	30	0.05	0.06
CEM	-4	16	0.28	0.28
INTI	-7	22	0.37	0.36
NIST	2	18	0.09	0.09
INMETRO1	-8	1	0.90	0.89

Median	**16.0 nm**
Simple arithmetic mean:	**17.4 nm**
standard uncertainty	4.2 nm
Variance weighted mean:	**10.0 nm**
standard uncertainty	0.6 nm
Variance weighted mean sans INMETRO1:	**17.0 nm**
standard uncertainty	4.4 nm
Observed chi-squared	**4.9**
Observed chi-squared sans INMETRO1	2.4
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\}$	0.56
$\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\}$ sans INMETRO1	0.88
Reduced chi-squared	**0.81**
sans INMETRO1	0.40
Birge Ratio	**0.90**
sans INMETRO1	0.63

2 mm Tungsten Carbide				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	1	31	0.05	0.05
CENAM	11	27	0.46	0.46
INMETRO2	3	30	0.12	0.13
CEM	11	16	0.72	0.71
INTI	-9	22	0.43	0.43
NIST	-12	18	0.68	0.67
INMETRO1	-7	5	0.75	0.74

Median	**$-$10.0 nm**
Simple arithmetic mean:	**$-$11.3 nm**
standard uncertainty	4.2 nm
Variance weighted mean:	**$-$16.6 nm**
standard uncertainty	1.9 nm
Variance weighted mean sans INMETRO1:	**$-$10.4 nm**
standard uncertainty	4.4 nm
Observed chi-squared	**7.4**
Observed chi-squared sans INMETRO1	5.0
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$	0.28
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ sans INMETRO1	0.55
Reduced chi-squared	**1.23**
sans INMETRO1	0.83
Birge Ratio	**1.11**
sans INMETRO1	0.91

100 mm Steel				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	-19	52	0.41	0.41
CENAM	12	45	0.29	0.28
INMETRO2	37	57	0.73	0.72
CEM	-43	33	1.33	1.31
INTI	1	41	0.02	0.02
NIST	5	35	0.15	0.14
INMETRO1	7	9	0.43	0.42

Median	**−100.0 nm**
Simple arithmetic mean:	**−104.9 nm**
standard uncertainty	8.0 nm
Variance weighted mean:	**−100.4 nm**
standard uncertainty	3.8 nm
Variance weighted mean sans INMETRO1:	**−111.9 nm**
standard uncertainty	8.6 nm
Observed chi-squared	**10.5**
Observed chi-squared sans INMETRO1	8.3
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\}$	0.11
$\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\}$ sans INMETRO1	0.22
Reduced chi-squared	**1.75**
sans INMETRO1	1.38
Birge Ratio	**1.32**
sans INMETRO1	1.17

50 mm Steel				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	0	37	0.01	0.01
CENAM	5	25	0.20	0.19
INMETRO2	27	38	0.79	0.79
CEM	-22	22	1.06	1.04
INTI	-12	28	0.48	0.47
NIST	5	26	0.19	0.19
INMETRO1	-1	6	0.11	0.11

Median	**31.0 nm**
Simple arithmetic mean:	**31.3 nm**
standard uncertainty	5.3 nm
Variance weighted mean:	**29.7 nm**
standard uncertainty	2.3 nm
Variance weighted mean sans INMETRO1:	**27.4 nm**
standard uncertainty	5.7 nm
Observed chi-squared	**6.8**
Observed chi-squared sans INMETRO1	6.6
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\rm obs}\}$	0.34
$\Pr\{\chi^2(\nu) > \chi^2_{\rm obs}\}$ sans INMETRO1	0.35
Reduced chi-squared	**1.14**
sans INMETRO1	1.11
Birge Ratio	**1.07**
sans INMETRO1	1.05

10 mm Steel				
	d_i	$k_{95}u(x_i)$	$\|E_{95,i}\| = d_i/k_{95}(d_i)u(d_i)$	$\|E_i\| = d_i/2u(d_i)$
NRC	14	31	0.52	0.57
CENAM	1	15	0.09	0.09
INMETRO2	30	30	1.14	1.22
CEM	-25	18	1.46	1.44
INTI	-7	22	0.33	0.33
NIST	-13	19	0.71	0.69
INMETRO1	-2	7	0.23	0.23

Median	**18.7 nm**
Simple arithmetic mean:	**20.7 nm**
standard uncertainty	3.9 nm
Variance weighted mean:	**18.5 nm**
standard uncertainty	1.5 nm
Variance weighted mean sans INMETRO1:	**16.7 nm**
standard uncertainty	4.1 nm
Observed chi-squared	**14.6**
Observed chi-squared sans INMETRO1	14.4
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$	0.02
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ sans INMETRO1	0.03
Reduced chi-squared	**2.43**
sans INMETRO1	2.40
Birge Ratio	**1.56**
sans INMETRO1	1.55

8 mm Steel				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	-20	31	0.73	0.80
CENAM	-4	14	0.28	0.28
INMETRO2	32	30	1.21	1.29
CEM	-2	16	0.13	0.13
INTI	-9	22	0.45	0.45
NIST	10	19	0.57	0.56
INMETRO1	-7	4	0.85	0.85

Median	**45.0 nm**
Simple arithmetic mean:	**49.0 nm**
standard uncertainty	3.8 nm
Variance weighted mean:	**43.1 nm**
standard uncertainty	1.5 nm
Variance weighted mean sans INMETRO1:	**48.8 nm**
standard uncertainty	3.9 nm
Observed chi-squared	**11.8**
Observed chi-squared sans INMETRO1	9.4
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\}$	0.07
$\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\}$ sans INMETRO1	0.15
Reduced chi-squared	**1.97**
sans INMETRO1	1.56
Birge Ratio	**1.41**
sans INMETRO1	1.25

5 mm Steel				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	1	31	0.05	0.05
CENAM	2	14	0.17	0.17
INMETRO2	22	30	0.84	0.90
CEM	-12	16	0.76	0.75
INTI	-12	22	0.58	0.58
NIST	1	18	0.08	0.08
INMETRO1	-4	7	0.51	0.51

Median	**-52.0 nm**
Simple arithmetic mean:	**-53.4 nm**
standard uncertainty	3.8 nm
Variance weighted mean:	**-57.1 nm**
standard uncertainty	1.5 nm
Variance weighted mean sans INMETRO1:	**-54.8 nm**
standard uncertainty	3.9 nm
Observed chi-squared	**6.2**
Observed chi-squared sans INMETRO1	5.8
Degrees of freedom:	6
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$	0.40
$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ sans INMETRO1	0.45
Reduced chi-squared	**1.04**
sans INMETRO1	0.97
Birge Ratio	**1.02**
sans INMETRO1	0.98

2 mm Steel				
	d_i	$k_{95}u(x_i)$	$\lvert E_{95,i}\rvert = d_i/k_{95}(d_i)u(d_i)$	$\lvert E_i\rvert = d_i/2u(d_i)$
NRC	-18	31	0.66	0.72
CENAM	1	14	0.07	0.07
INMETRO2	24	30	0.90	0.97
CEM	-10	16	0.65	0.64
INTI	-3	22	0.15	0.15
NIST	6	18	0.36	0.35
INMETRO1	0	7	0.01	0.01
			Median	**35.9 nm**
			Simple arithmetic mean:	**36.0 nm**
			standard uncertainty	3.8 nm
			Variance weighted mean:	**35.8 nm**
			standard uncertainty	1.5 nm
			Variance weighted mean sans INMETRO1:	**35.1 nm**
			standard uncertainty	3.9 nm
			Observed chi-squared	**6.7**
			Observed chi-squared sans INMETRO1	6.6
			Degrees of freedom:	6
			$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$	0.35
			$\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ sans INMETRO1	0.36
			Reduced chi-squared	**1.11**
			sans INMETRO1	1.11
			Birge Ratio	**1.05**
			sans INMETRO1	1.05

Table 12: List of median, simple arithmetic mean, variance weighted mean and statistical consistency parameters chi-squared, Birge Ratio and normalized deviations $|E_i|$ and $|E_{95,i}|$.

A.2 Discussion

Consistency tests of chi-squared, Birge ratio and $\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\}$ probablility are evaluated based on the inverse variance weighted mean, keeping in mind that these tests are designed to test statistical consistency only in the case of the variance weighted mean. Values for these parameters are listed in Tables below and plotted in Figure 7. For the case of mean value evaluated with equal weights (arithmetic mean), sophisticated consistency indicators must be evaluated by other methods. This Appendix reports on the simple comparison of d_i with the participant's claimed $U_{95,i}$.

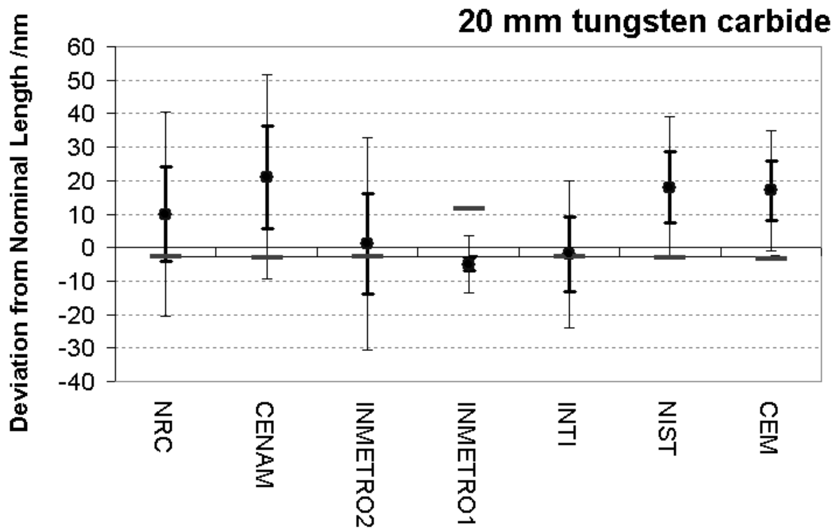
The 10 mm steel gauge block data set could be considered discrepant. At the time of writing this report, results from pilot measurements over the time duration of Stage Two of the comparison show that this gauge block was shrinking during the time of this comparison. However at the time when Stage One was just completed, it was not obvious from the pilot measurements taken during the slice of time duration for Stage One of the comparison that the gauge block was indeed shrinking. For this reason, results from the 10 mm steel gauge block are left out of the discussions of statistical consistency.

Now the technique of *exclusive statistics* [10] provides a simple and statistically rigorous procedure for demonstrating the amount of correlation of each lab with the evaluated mean. The 'exclusive mean' is the mean value (arithmetic or otherwise) evaluated for each participant in turn, omitting the participants own value from the calculation. The exclusive mean value includes the values submitted by each of the other participants, but excludes the value of the 'exclusive mean participant'. The exclusive mean expresses the results of a comparison from the point of view of how each laboratory performs with respect to the rest of the world. The plots of the data in the body of the report (Figures 1 and 2) show the simple arithmetic mean evaluated from all participants as a solid line. The exclusive simple arithmetic mean is plotted as the short thick lines for each participant. The difference between the inclusive and exclusive means allows us to graphically observe correlations between individual labs and the mean value. For the simple arithmetic mean case, it is clear that none of the individual labs have a dominating influence on the mean value. However, one can observe strong correlation between INMETRO1 and the inverse variance weighted mean because of the relatively small measurement uncertainties claimed by INMETRO1. The 5 mm tungsten carbide and 20 mm tungsten carbide gauges are the worst-case examples; they are shown in Figure 6. In general, INMETRO1 results dominate the inverse variance weighted mean by a relative amount ranging from 78 % to 98 %.

Reduced chi-squared, Birge Ratio and $\Pr\{\chi^2(\nu) > \chi^2_{\mathrm{obs}}\}$ are evaluated both including all participants and a second time excluding the results of INMETRO1 for the reason that INMETRO1 submitted very optimistic uncertainty claims relative to conventional capabilities[3]. For most gauge blocks, these consistency tests are either improved or remain the same when INMETRO1 results are included (see below). The measurement values themselves cannot be considered to be outliers since inclusion or exclusion of x_{INMETRO1} for any of the gauge blocks of the comparison does not change the arithmetic mean or median values significantly. This comparison analysis at once probes the impact of the small standard uncertainties and low degrees of freedom submitted by INMETRO1 on the KCRV, and attempts to settle on a KCRV that represents all participants fairly.

The Tables list the simple consistency indicator of comparing d_i evaluated with the simple arithmetic mean value with each participants submitted U_{95}. Examination of $d_i < k_{95}u(x_i)$ largely passes for all participants for all gauge blocks with few discrepancies considering the expected 5 % based on statistics. However, $d_i < k_{95}u(x_i)$ fails for 6 out of the 11 gauge blocks for INMETRO1, yet in all except one case INMETRO1 passes the traditional normalized deviation $|E_i|$ and $|E_{95,i}|$. Interestingly, the extra precision offered by the evaluation of $|E_{95,i}|$ does not result in appreciable differences from the result of the approximate $|E_i|$.

[3]The technical debate as to the validity of INMETRO1 technique and uncertainty analysis is deferred to discussion in the literature.



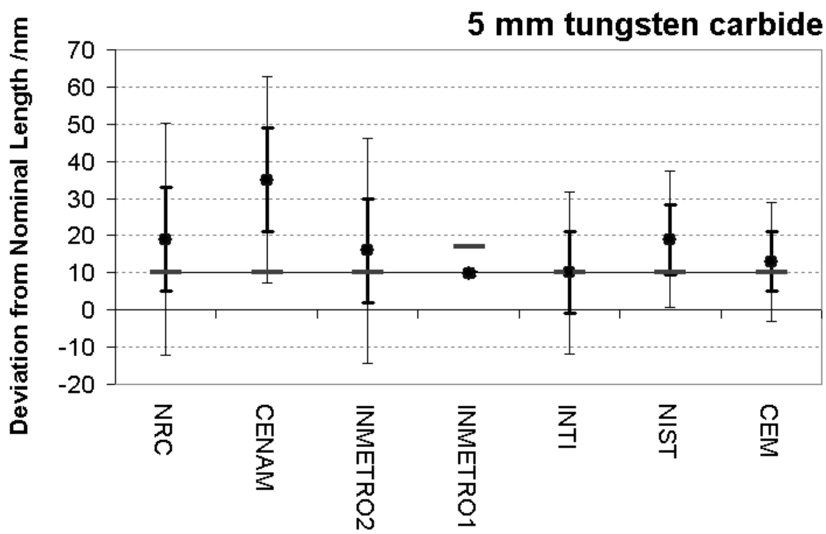


Figure 6: Plots of central length expressed as deviation from nominal length reported by each participant for 20 mm and 5 mm tungsten carbide gauge blocks. Thick error bars represent the standard uncertainty, while longer thin error bars represent $k_{95}u(x_i)$ where $k_{95} = t_p(\nu_i)$ from the Student's t-distribution for standard uncertainties $u(x_i)$ and degrees of freedom ν_i submitted by the participants. The solid line represents the inverse variance weighted mean evaluated taking results of all participants. The dash for each participant represents the exclusive inverse variance weighted mean (see text).

For most steel gauge blocks, the chi-squared and Birge Ratio values remain much the same whether or not INMETRO1 results are included. Although the 100 mm steel and the 8 mm steel gauge blocks are on the verge of discrepant, they are somewhat improved with the exclusion of INMETRO1 results. Values of $\Pr\{\chi^2(\nu) > \chi^2_{\text{obs}}\}$ are generally improved with INMETRO1 excluded.

The results for tungsten carbide gauge blocks appear to be more sensitive to the weighting influence of INMETRO1. In general, larger differences between weighted mean and simple arithmetic mean and the consistency parameters is observed. For the two gauge blocks in Figure 6, INMETRO1 results could be considered as outliers relative to the variance weighted mean evaluated for the sub-set of the participants excluding INMETRO1 (the exclusive variance weighted mean). The 20 mm tungsten carbide results could be considered discrepant; and this gauge block was not changing length during the comparison. The INMETRO1 result for this gauge is 4 standard deviations away from the exclusive variance weighted mean value, and the chi-squared, Birge ratio and probablility results reflect that the data sub-set would be more consistent than the full data set containing all participants. For the 5 mm gauge block the INMETRO1 result is about 11 standard deviations away from the exclusive variance weighted mean, and the reduced chi-squared and Birge ratio parameters indicate that the full data set including all participants is more consistent. This is because the INMETRO1 result with the very small measurement uncertainty 'owns' the variance weighted mean, and is therefore very consistent with it.

The weighted mean is not recommended as the KCRV when not all self-declared uncertainties are considered reliable [17]. The very low degrees of freedom reported by INMETRO1, and the dominating weight of their results to the variance weighted mean as demonstrated by exclusive statistics and the behaviour of the consistency parameters, provide reasonable evidence to select the simple arithmetic mean for the SIM.4.2 Stage One KCRV.

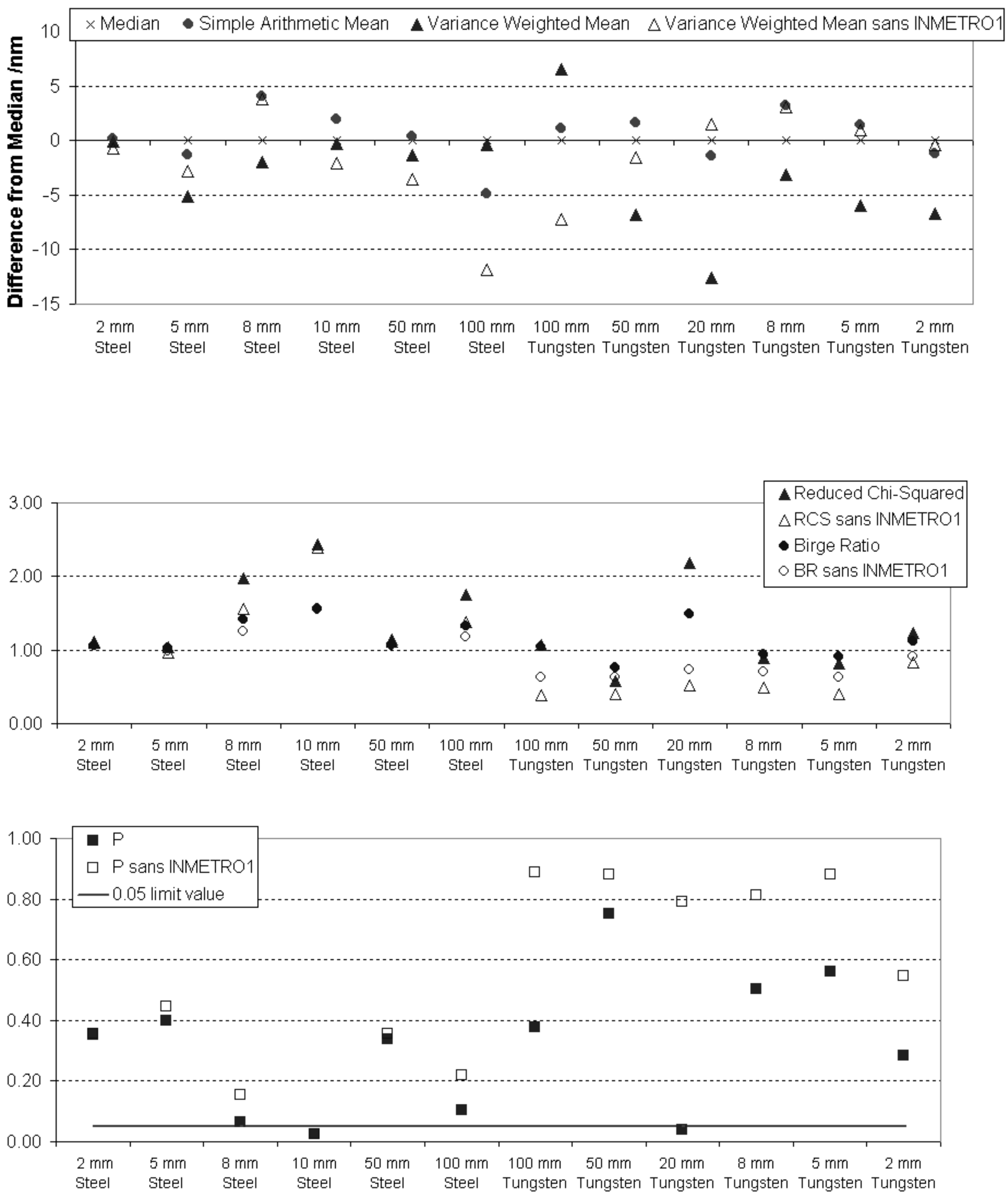


Figure 7: Graphical representation of mean values and consistency parameters evaluated for the comparison data.

Participant	Nominal Length of Steel Gauge Blocks /mm					
	2	5	8	10	50	100
NRC	-18 ± 25	1 ± 25	-20 ± 25	14 ± 25	0 ± 32	-19 ± 47
CENAM	1 ± 14	2 ± 14	-4 ± 14	1 ± 15	5 ± 24	12 ± 42
INMETRO2	24 ± 25	22 ± 25	32 ± 25	30 ± 25	27 ± 34	37 ± 52
INMETRO1	0 ± 8	-4 ± 8	-7 ± 8	-2 ± 8	-1 ± 11	7 ± 18
INTI	-3 ± 20	-12 ± 20	-9 ± 20	-7 ± 20	-12 ± 26	1 ± 39
NIST	6 ± 17	1 ± 18	10 ± 18	-13 ± 18	5 ± 25	5 ± 34
CEM	-10 ± 16	-12 ± 16	-2 ± 16	-25 ± 17	-22 ± 21	-43 ± 33

Table 13: Difference between participant value and the KCRV (simple arithmetic mean) listed with the $k = 2$ expanded uncertainty $U(d_i)$ for steel gauge blocks.

Participant	Nominal Length of Tungsten Carbide Gauge Blocks /mm					
	2	5	8	20	50	100
NRC	1 ± 25	2 ± 25	8 ± 25	1 ± 25	5 ± 31	-13 ± 44
CENAM	11 ± 25	18 ± 25	17 ± 25	12 ± 28	17 ± 38	18 ± 63
INMETRO2	3 ± 25	-1 ± 25	-3 ± 25	-8 ± 27	6 ± 32	17 ± 47
INMETRO1	-7 ± 9	-8 ± 8	-8 ± 9	-14 ± 10	-10 ± 12	9 ± 18
INTI	-9 ± 20	-7 ± 20	-9 ± 20	-11 ± 21	-6 ± 23	-12 ± 32
NIST	-12 ± 18	2 ± 18	-6 ± 18	9 ± 20	-2 ± 25	-1 ± 34
CEM	11 ± 16	-4 ± 16	2 ± 16	8 ± 18	-13 ± 20	-18 ± 27

Table 14: Difference between the participant value and the KCRV (simple arithmetic mean) listed with the $k = 2$ expanded uncertainty $U(d_i)$ for tungsten carbide gauge blocks.

B Degrees of Equivalence

Tables of degrees of equivalence list the difference between the measurement value submitted by each participant and the KCRV as described by equation (6) with $x_{\mathrm{ref}} \equiv \overline{x}$. The expanded uncertainty $U(d_i) = 2u(d_i)$ where $u(d_i)$ is calculated from equation (11) for the $n = 7$ participants of the comparison. On the basis of statistical variability alone, 5 % of measurements would be expected to be classified as discrepant. Values in Tables 13 and 14, and the above discussion allude to the suggestion that most participant labs could have been conservative in their estimate of uncertainties.

C Linking of SIM.4.2 Gauge Block Comparison to the CCL-K1 Gauge Block Comparison

At its 11th meeting in 2003, the Consultative Committee for Length (CCL) decided that "artefact-based key comparisons in dimensional metrology will not use a numerical link between a CCL key comparison and any corresponding RMO comparison. Instead, the link will be based on competencies demonstrated by the participant laboratory which took part as linking NMIs in the CCL and RMO key comparisons. The CCL and RMO key comparisons will be deemed as being equivalent." If the linking NMIs were judged to have performed competently in both comparisons (CCL, RMO), then the comparisons were to be regarded as equivalent. The judgment of the competence is the responsibility of the WGDM upon consideration of the Draft B report. The Sistema Interamericano de Metrología (SIM) Regional Comparison of gauge block calibration by interferometry SIM.4.2 links to the CCL Key Comparison CCL-K1 [21] through the participation of the following national metrology institutes: National Research Council Canada Institute for National Measurement Standards (NRC-INMS), the United States National Institute of Standards and Technology (NIST) and the Centro Nacional de Metrología (CENAM) of Mexico.

D Tables of Bilateral Equivalence

The degree of equivalence between institute i and institute j is listed as a pair of values where

$$d_{i,j} = x_i - x_j \tag{14}$$

and

$$U_{95}(d_{i,j}) = k_{95}\ u(d_{i,j}) \tag{15}$$

where $u^2(d_{i,j}) = u^2(x_i) + u^2(x_j)$, and k_{95} is evaluated from the Student's t-distribution and the effective degrees of freedom ν_{eff} determined by the Welch-Satterthwaite approximation from ν_i and ν_j submitted by the participants according to Section G of the GUM [22].

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-19 ± 33						
INMETRO2	-42 ± 41	-23 ± 33					
INMETRO1	-18 ± 31	1 ± 14	24 ± 30				
INTI	-15 ± 37	4 ± 26	27 ± 36	3 ± 22			
NIST	-24 ± 35	-5 ± 23	18 ± 34	-6 ± 18	-9 ± 28		
CEM	-8 ± 34	11 ± 21	34 ± 33	10 ± 16	7 ± 27	16 ± 24	

Table 15: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 2 mm steel gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-1 ± 33						
INMETRO2	-21 ± 41	-20 ± 33					
INMETRO1	6 ± 31	7 ± 14	27 ± 30				
INTI	13 ± 37	14 ± 26	34 ± 36	8 ± 22			
NIST	0 ± 35	1 ± 23	21 ± 35	-6 ± 19	-13 ± 29		
CEM	13 ± 34	14 ± 21	34 ± 33	8 ± 16	0 ± 27	13 ± 24	

Table 16: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 5 mm steel gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-16 ± 33						
INMETRO2	-52 ± 41	-36 ± 33					
INMETRO1	-13 ± 31	3 ± 15	39 ± 30				
INTI	-11 ± 37	5 ± 26	41 ± 36	2 ± 22			
NIST	-30 ± 35	-14 ± 24	22 ± 35	-17 ± 19	-19 ± 29		
CEM	-18 ± 34	-2 ± 21	34 ± 33	-5 ± 16	-7 ± 27	12 ± 25	

Table 17: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 8 mm steel gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	13 ± 33						
INMETRO2	-16 ± 41	-29 ± 33					
INMETRO1	16 ± 31	3 ± 15	32 ± 30				
INTI	21 ± 37	8 ± 26	37 ± 36	5 ± 22			
NIST	27 ± 35	14 ± 24	43 ± 35	11 ± 19	6 ± 29		
CEM	39 ± 35	26 ± 23	55 ± 34	23 ± 18	18 ± 28	12 ± 26	

Table 18: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 10 mm steel gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-5 ± 47						
INMETRO2	-27 ± 52	-22 ± 46					
INMETRO1	1 ± 37	6 ± 26	28 ± 39				
INTI	12 ± 45	17 ± 37	39 ± 47	11 ± 28			
NIST	-5 ± 45	0 ± 36	22 ± 46	-6 ± 27	-17 ± 38		
CEM	22 ± 42	27 ± 33	49 ± 44	21 ± 22	10 ± 35	27 ± 34	

Table 19: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 50 mm steel gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-31 ± 73						
INMETRO2	-56 ± 77	-25 ± 72					
INMETRO1	-27 ± 52	5 ± 46	30 ± 58				
INTI	-20 ± 66	11 ± 61	36 ± 70	7 ± 42			
NIST	-24 ± 62	7 ± 57	32 ± 67	3 ± 36	-4 ± 54		
CEM	24 ± 61	55 ± 56	80 ± 66	51 ± 34	44 ± 53	48 ± 48	

Table 20: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 100 mm steel gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-10 ± 42						
INMETRO2	-2 ± 41	8 ± 40					
INMETRO1	8 ± 31	18 ± 28	10 ± 31				
INTI	10 ± 37	20 ± 35	12 ± 36	2 ± 22			
NIST	13 ± 35	23 ± 33	15 ± 34	5 ± 18	3 ± 28		
CEM	-10 ± 34	0 ± 32	-8 ± 33	-18 ± 16	-20 ± 27	-23 ± 24	

Table 21: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 2 mm tungsten carbide gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-16 ± 42						
INMETRO2	3 ± 41	19 ± 40					
INMETRO1	9 ± 31	25 ± 28	6 ± 30				
INTI	9 ± 37	25 ± 35	6 ± 36	0 ± 22			
NIST	0 ± 35	16 ± 33	-3 ± 35	-9 ± 18	-9 ± 29		
CEM	6 ± 34	22 ± 32	3 ± 33	-3 ± 16	-3 ± 27	6 ± 24	

Table 22: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 5 mm tungsten carbide gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-9 ± 42						
INMETRO2	11 ± 41	20 ± 40					
INMETRO1	16 ± 31	25 ± 28	5 ± 31				
INTI	17 ± 37	26 ± 35	6 ± 36	2 ± 22			
NIST	14 ± 35	23 ± 33	3 ± 35	-2 ± 19	-3 ± 29		
CEM	6 ± 34	15 ± 32	-5 ± 33	-10 ± 16	-11 ± 27	-8 ± 25	

Table 23: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 8 mm tungsten carbide gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-11 ± 44						
INMETRO2	9 ± 42	20 ± 43					
INMETRO1	15 ± 31	26 ± 31	6 ± 32				
INTI	12 ± 36	23 ± 37	3 ± 38	-3 ± 22			
NIST	-8 ± 36	3 ± 37	-17 ± 37	-23 ± 21	-20 ± 30		
CEM	-7 ± 34	4 ± 35	-16 ± 36	-22 ± 18	-19 ± 28	1 ± 27	

Table 24: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 20 mm tungsten carbide gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-12 ± 58						
INMETRO2	-1 ± 50	11 ± 55					
INMETRO1	15 ± 35	27 ± 43	16 ± 37				
INTI	11 ± 42	23 ± 49	12 ± 43	-4 ± 24			
NIST	7 ± 43	19 ± 50	8 ± 45	-8 ± 27	-4 ± 35		
CEM	18 ± 40	30 ± 47	19 ± 41	3 ± 21	7 ± 31	11 ± 33	

Table 25: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 50 mm tungsten carbide gauge block.

	NRC	CENAM	INMETRO2	INMETRO1	INTI	NIST	CEM
NRC							
CENAM	-31 ± 91						
INMETRO2	-30 ± 70	1 ± 87					
INMETRO1	-22 ± 49	9 ± 71	8 ± 52				
INTI	-1 ± 57	30 ± 77	29 ± 60	21 ± 32			
NIST	-12 ± 59	19 ± 79	18 ± 62	10 ± 36	-11 ± 47		
CEM	5 ± 54	36 ± 75	35 ± 57	27 ± 27	6 ± 40	17 ± 43	

Table 26: Bilateral equivalence $d_{i,j} \pm k_{95}u(d_{i,j})$ for 100 mm tungsten carbide gauge block.