CCL Key Comparison

CCL-K1:2011

Calibration of gauge blocks by Optical Interferometry

Comparison Report

Final Report

Miguel Viliesid,
Carlos Colín
Centro Nacional de Metrología, CENAM

Pierre Dubé
National Research Council Canada, NRC-CNRC

July 2022

CONTENT

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 the Working Group on Dimensional Metrology (WGDM[1]) of the Consultative Committee for Length (CCL) meeting held at INRiM in Torino, Italy, in 2008 it was decided to start a new Key Comparison (KC) on Gauge Blocks (GB) calibration by optical interferometry. It was decided to merge the short GB comparison CCL-K1 (0,5 mm to 100 mm) with the long GB comparison (over 100 mm to 500 mm), previously named CCL-K2. The designated pilot laboratories were CENAM from Mexico for short GB and NRC-CNRC[2] from Canada for long GB. The present report is the Draft B version. It contains measurement results as well as preliminary data analysis. It is addressed only to the participants and must be kept confidential at present.

A measurement protocol was drawn consistent with the corresponding GB comparisons of CCL and participants were requested to follow it as strictly as possible. Due to the large number of participants, the participating National Metrology Institutes (NMIs) were strongly advised to adjust to the allocated dates to perform their measurements.

In order to have comparable uncertainties from the participants, it was suggested to estimate them in a similar way, therefore, the protocol recommended to apply a proposed model taken from [3] with its corresponding uncertainty budget. Only in those cases where the calibration method was different, was a different model applied.

By their declared intention to participate in this Key Comparison, the participating laboratories accepted the general instructions of the protocol.

The present report was prepared by Dr. Pierre Dubé from NRC-CNRC and Eng. Carlos Colín and Dr. Miguel Viliesid from CENAM.

The results of this international comparison contribute to the Mutual Recognition Arrangement (MRA) between the NMIs of the Metre Convention. This CIPM key comparison should be combined with Regional Comparisons (Regional Metrology Organizations KC) following exactly the same scheme, with laboratories participating in both the CIPM and the RMO KC in order to be able to establish equivalence among all participating NMIs in different regions.

2. Organization

The pilot laboratories representatives Dr. Pierre Dubé from NRC-CNRC for long gauge blocks, and Dr. Miguel Viliesid and Eng. Carlos Colín from CENAM for short gauge blocks, coordinated the comparison exercise.

[1] Former name. At present it corresponds to WG-MRA, Working Group in Mutual Recognition Arrangement.
[2] Formerly NRC-INMS, Institute for National Research Standards.

2.1 Requirements for participation

According to the WGDM recommendation No. 2 (document CCDM/WGDM/97-50b), the participating laboratories should offer these calibration services regularly to their customers. The participant NMIs should declare a measurement uncertainty (at k=1) below a required uncertainty level. These uncertainty levels were reviewed during the meeting of the CCL Sub-Group on Key Comparisons (sWG-KC) in June 2010 (Singapore) and it was agreed to fix the standard uncertainties as:

- $u = 0{,}02\ \mu\text{m} + 1{,}3 \times 10^{-7}\ L$, for short gauge blocks and
- $u = 0{,}02\ \mu\text{m} + 1{,}6 \times 10^{-7}\ L$, for long gauge blocks

Adjusting by regression these limiting uncertainties to the usual quadratic notation[3] on the corresponding intervals, 0 mm to 100 mm for short GB and over 100 mm for long GB, and expressing it in the usual units posted for the CMCs in the BIPM webpage we obtain that it corresponds approximately to:

- $u = \text{Q}[21\ \text{nm}, 2.6\text{x}10^{-7}\ L]$, for short gauge blocks and
- $u = \text{Q}[32\ \text{nm}, 1.8\text{x}10^{-7}\ L]$, for long gauge blocks

It was recommended that at least two laboratories from each metrology region should participate in this Key Comparison in order to be able to link to the corresponding regional comparisons.

2.2 Participants

There were twelve participants. Three from APMP, three from EURAMET, one from COOMET[4], four from SIM and one from AFRIMETS. The list of participants and their contact information are given in **Table 1**.

2.3 Original Time schedule

Participants were assigned six weeks to perform their measurements, which included customs clearance and transportation to the next participant. The participants were asked to strictly respect the time schedule. The pilot laboratories carried out one intermediate measurement check during the circulation. The original allocated dates appear in **Table 2**.

2.4 Actual Time schedule

Table 3 shows the actual circulation dates and dates of reception of results by the pilot laboratories. The circulation scheme of the artifacts was delayed at the end of the first circulation loop due to customs problems in Brazil. Otherwise, Japan asked to be allocated a new time slot at the end of the second loop because they were experiencing power shortages in the aftermath of the earthquake and tsunami catastrophe of March 2011. As a consequence, the second loop was rescheduled as shown.

[3] $\text{u} = \text{Q}[\text{a}, \text{bL}] = \sqrt{\text{a}^2 + \text{b}^2 \cdot \text{L}^2}$

[4] Although PTB is part of both EURAMET and COOMET, it participated as part of COOMET for the present Key Comparison.

CCL-K1:2011 Participants					
Country	**Institute**	**Contact person**	**Postal Address**	**Tel. / FAX**	**e-mail**
Brazil	**INMETRO**	Hans Peter H. Grieneisen, Luiz V. Tarelho, & Ricardo dos S. França	Av. Nossa Senhora de Graças, 50 Xerém, Duque de Caxias, RJ, BRASIL, CEP 25250-020	(21) 2679-9077 / Fax (21) 2679-1507	hpgrieneisen@inmetro.gov.br; lvtarelho@inmetro.gov.br rsfranca@inmetro.gov.br
Canada	**NRC-CNRC**	Pierre Dubé	1200 Montreal Road Building M-36 Room 146 Ottawa, Ontario, CANADA K1A 0R6	+1.613.998-6768 / Fax +1.613.952.1394	pierre.dube@nrc-cnrc.gc.ca
China	**NIM**	Sitian Gao	No.18, Bei San Huan Dong Lu, Chaoyang Dist, Beijing, P. R. of CHINA 100013	+86-10-64218565 / Fax +86-10-64218703	Gaost@nim.ac.cn
Finland	**MIKES**	Antti Lassila	Tekniikantie 1 FI-02151 Espoo, P. O. Box 9, FINLAND	+358 10 6054 000 / Fax +358 10 6054 499	antti.lassila@mikes.fi
Germany	**PTB**	Harald Bosse	Abteilung 5: Fertigungsmess-technik (Div. 5: Prec. Eng.) Bundesallee 100 38116 Braunschweig, GERMANY	+49 531 592 5010 Fax +49 531 592 5015	Harald.Bosse@ptb.de
Italy	**INRIM**	Alessandro Balsamo & Paola Pedone	Strada delle Cacce 91 IT-10135 Torino, ITALY	+39 011 3919 970 / Fax: +39 011 3919 959	a.balsamo@inrim.it p.pedone@inrim.it
Japan	**NMIJ-AIST**	Toshiyuki Takatsuji	AIST Central 3, 1-1-1 Umezono, Tsukuba, Ibaraki 305-8563, JAPAN	+81-29-861-4361 / Fax +81-29-861-4152	toshiyuki.takatsuji@aist.go.jp
Mexico	**CENAM**	Miguel Viliesid & Carlos Colín	km 4,5 Carretera Los Cués, El Marqués, 76246, Querétaro., MEXICO	+52-442 2110574 / Fax +52-442 2110579	miguel.viliesid@cenam.mx; ccolin@cenam.mx
Singa-pore	**A*STAR-NMC**	Siew-Leng Tan	2 Fusionopolis Way, #08-05, Innovis, SINGAPORE 138634	+65 67149256	tan_siew_leng@nmc.a-star.edu.sg
South Africa	**NMISA**	Oelof Kruger	Private Bag X34 Lynnwood Ridge, Pretoria 0040, SOUTH AFRICA	+27 12 841 4152 / Fax +27 12 841 2131	oakruger@nmisa.org
Switzer-land	**METAS**	Rudolf Thalmann	Lindenweg 50 CH-3003 Bern-Wabern, SWITZERLAND	+41 31 32 33 385 / Fax +41 31 32 33 210	rudolf.thalmann@metas.ch
United States	**NIST**	Jack Stone & John Stoup	100 Bureau Drive, M/S 8210 Gaithersburg, Maryland 20899-8210, USA	301-975-5638 / Fax 301-869-0822	jack.stone@nist.gov john.stoup@nist.gov

Table 1. Participating NMIs and contact persons.

During the second loop, there were further delays in Singapore and South Africa. Because of these delays, the ATA Carnet was about to expire and the GBs had to be sent back to Canada before they could be sent to Japan, the last participant. Finally, the artifacts came back to the pilot laboratories but then another complication happened at CENAM. We experienced problems with the GB interferometer from January 2013 until March 2014. This hindered us to make the closure measurements as well as the long GB measurements. We give a detailed explanation in **Appendix A**.

CCL-K1:2011 Original Time Schedule				
Region	**Laboratory**	**City**	**Country**	**Date of Reception**
Pilot Laboratory Short GB	**CENAM**	Querétaro	MEXICO	2011-03-14
Pilot Laboratory Long GB	**NRC-CNRC**	Ottawa	CANADA	2011-05-01
EURAMET	**METAS**	Bern-Wabern	SWITZERLAND	2011-07-04
	MIKES	Espoo	FINLAND	2011-08-18
	INRIM	Torino	ITALY	2011-10-02
COOMET	**PTB**	Braunschweig	GERMANY	2011-11-16
SIM	**NIST**	Gaithersburg	USA	2011-12-31
	INMETRO	Xerém	BRAZIL	2012-02-14
Pilot Laboratory Long GB	**NRC-CNRC**	Ottawa	CANADA	2012-03-30
Pilot Laboratory Short GB	**CENAM**	Querétaro	MEXICO	2012-04-06
APMP	**NMIJ-AIST**	Tsukuba	JAPAN	2012-05-14
	NIM	Beijing	P. R. of CHINA	2012-06-28
	A*STAR-NMC	Singapore	SINGAPORE	2012-08-12
AFRIMETS	**NMISA**	Pretoria	SOUTH AFRICA	2012-09-26
Pilot Laboratory Long GB	**NRC-CNRC**	Ottawa	CANADA	2012-11-10
Pilot Laboratory Short GB	**CENAM**	Querétaro	MEXICO	2012-12-25

Table 2. Original scheduled dates of measurement.

Participant laboratories were given four weeks after completion of the GB measurements to submit their measurement results to the pilot laboratories. However, only PTB, INMETRO, and NMIJ-AIST respected this period.

2.5 Transportation.

The short GBs were packed in a wooden case as shown in **Figure 1** with slots of polystyrene foam of the size of each GB to avoid any damage. A data logger was also included inside the casing of the short GBs to register the temperature variations throughout the circulation.

The long GBs were also packed separately in a wooden case with compartments of the size of each GB to prevent damage. The wooden cases of the short and long GBs were packaged into a robust hard plastic case and extra packaging foam for shipping.

CCL-K1:2011 Actual Time Schedule					
Region	**Laboratory**	**Country**	**Date of Reception of GB**	**Date of Dispatch of GB**	**Date of Reception of Results**
Pilot Laboratory Short GB	**CENAM**	MEXICO	- - -	2011-04-15	2011-07-22[5]
Pilot Laboratory Long GB	**NRC-CNRC**	CANADA	2011-05-15[6]		2011-08-11
EURAMET	**METAS**	SWITZERLAND	2011-07-04	2011-07-19	2011-12-26
	MIKES	FINLAND	2011-08-03	2011-09-02	2011-12-16
	INRIM	ITALY	2011-09-06	2011-11-07	2013-06-04
COOMET	**PTB**	GERMANY	2011-11-10	2011-12-21	2012-01-24
SIM	**NIST**	USA	2012-01-10	2012-02-24	2014-09-02
	INMETRO	BRAZIL	2012-03-21	2012-05-26	2012-05-21
Pilot Laboratory Long GB[7]	**NRC-CNRC**	CANADA	2012-06-15	2012-08-01	
Pilot Laboratory Short GB[6]	**CENAM**	MEXICO	2012-06-28	2012-07-04	
APMP	**NIM**	P. R. of CHINA	2012-09-04	2012-09-27	2013-05-27
	A*STAR-NMC	SINGAPORE	2012-10-09	2012-12-07	2013-04-11
AFRIMETS	**NMISA**	SOUTH AFRICA	2013-01-15	2013-05-13[8]	2013-09-23
APMP	**NMIJ-AIST**	JAPAN	2013-06-03	2013-07-11	2013-08-05
Pilot Laboratory Long GB	**NRC-CNRC**	CANADA	2013-07-18	2013-08-08	
Pilot Laboratory Short GB	**CENAM**	MEXICO	2013-08-29	- - -	2014-03-16[9]

Table 3. Actual dates of measurement and reception of results.

[5] Submitted to Andrew Lewis from NPL who accepted to serve as escrow for the results of the pilot laboratories.
[6] Twelve short GBs were sent on 2011-04-28 and the 0.5 mm block was sent afterwards to NRC-CNRC on the indicated date.
[7] NRC-CNRC shipped and received short and long GB to and from the participants and then sent the short GBs to CENAM for control measurements. CENAM sent them back to NRC-CNRC at the indicated date to then dispatch all the GBs to the following participant.
[8] Sent back to the copilot NRC-CNRC to have the ATA Carnet renewed.
[9] Long GB results of CENAM submitted to Long GB copilot NRC-CNRC. As explained in Appendix A, CENAM experienced problems with its GB interferometer. The problems were spotted in July 2013 and not solved until March 2014. This was the reason for the delay in the submission of results.

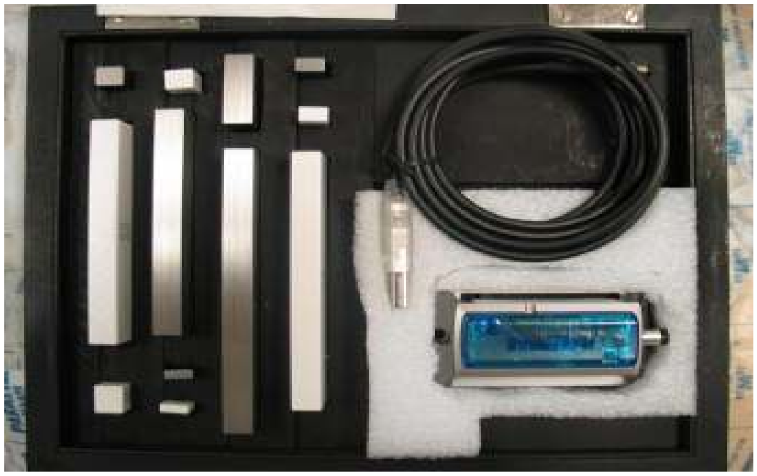

Figure1. Short GB inside case.

Items included in the circulation were:

Case for short GBs containing the following items:

- 7 steel short gauge blocks
- 6 ceramic short gauge blocks
- Temperature data logger with USB cable

Case for long GBs containing the following items:

- 3 steel long gauge blocks

Along with:

- Handling instructions
- 1 copy of the measurement protocol
- A pair of cotton gloves
- A pair of plastic gloves
- Plastic case containing the wooden cases

Outside the plastic case was attached a sleeve accessible to the customs officer containing:

- The ATA Carnet

Participants were instructed to protect the steel GBs with appropriate oil before shipping to the next participant and to use a reliable parcel service knowledgeable on ATA Carnet import export of goods.

Upon reception, participants were asked to inform the pilot laboratories by sending out to the pilot laboratories **Annex A1: Reception of Standards**; and once shipped they had to inform the pilot laboratories by filling **Annex A2**: **Shipment of Standards** indicating all pertinent information. These Annexes were included in the measurement protocol document.

2.6 Financial aspects and insurance.

Each participating laboratory had to cover the costs of shipping, customs formalities and transport insurance against loss or damage. The organization costs were covered by the pilot laboratories, which included the length standards, the data logger, the cases and packaging, and the ATA Carnet.

3. Description of the circulated standards

Thirteen short gauge blocks, seven made of steel and six of ceramics, and three steel long gauge blocks were selected for the comparison exercise. The GBs had a rectangular cross section and complied with the ISO 3650 Standard [1]. The nominal lengths, identification numbers, coefficients of thermal expansion (CTEs) and uncertainties ($k = 2$), and the names of manufacturers are given in **Tables 4**, **5** and **6**.

Nominal Length, mm	Identification Number	CTE (10^{-6} K^{-1})	Manufacturer
0.5	000774	10.9 ± 1.0	Mitutoyo
3	090771	10.8 ± 0.5	Mitutoyo
5	090772	10.8 ± 0.5	Mitutoyo
7	091740	10.8 ± 0.5	Mitutoyo
25	053292	10.8 ± 0.5	Mitutoyo
80	081305	10.8 ± 0.5	Mitutoyo
100	053178	10.8 ± 0.5	Mitutoyo

Table 4. Steel short gauge blocks

Nominal Length, mm	Identification Number	CTE (10^{-6} K^{-1})	Manufacturer
3	090486	9.3 ± 0.5	Mitutoyo
5	090662	9.3 ± 0.5	Mitutoyo
7	090343	9.3 ± 0.5	Mitutoyo
10	090721	9.3 ± 0.5	Mitutoyo
80	080099	9.3 ± 0.5	Mitutoyo
90	080173	9.3 ± 0.5	Mitutoyo

Table 5. Ceramic short gauge blocks

Nominal Length, mm (inches[10])	Identification Number	CTE (10^{-6} K^{-1})	Manufacturer
152.4 (6)	36686	10.5 ± 0.6	Hommel Werke
254.0 (10)	36686	10.5 ± 0.6	Hommel Werke
508.0 (20)	36686	10.5 ± 0.6	Hommel Werke

Table 6. Steel long gauge blocks

4. Inspection and reported condition of gauge blocks upon reception

4.1 Condition of the gauge blocks

The GBs had to be measured based on the standard procedure that the participants regularly use for this calibration service for their customers. Before making the measurements, the GBs were to be inspected by the recipient NMI to verify that the measuring surfaces were not damaged and that they did not present severe scratches, rust, or burrs that could affect or hinder the measurements. Participants had to fill **Annexes B1** and **B2 of the protocol** to state the conditions of the GBs as received. The designation of identification for the measuring faces was as shown on **Figures 2** and **3**.

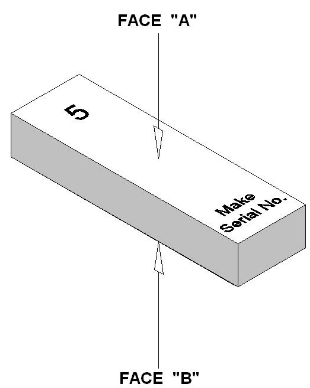

Figure 2. Gauge Block with nominal length < 6 mm.

[10] The long gauge blocks chosen for this key comparison have nominal lengths in imperial units because of technical reasons regarding their quality and quality of wringing. This choice is not an indication that imperial units are recommended. The nominal lengths and their deviations are stated in SI units in the current report.

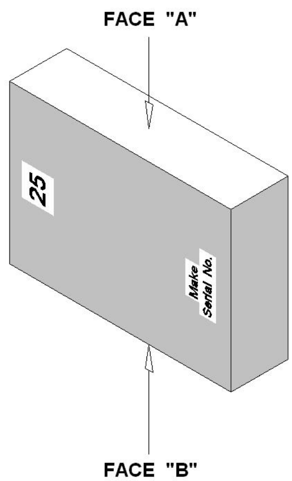

Figure 3. Gauge block with nominal length ≥ 6 mm.

The whole set of circulated GBs made the complete two-loop tour and came back to the pilot laboratories in good enough condition to be measured by the pilot laboratories. However, some of the GBs suffered some damage during circulation other than normal slight damage from the wringing process.

The first participant, METAS, reported scratches and damage on almost all of the steel GBs and that the 5 mm GB was difficult to wring[11]. Fortunately, it turned out that the scratches were slight and typical of GBs that are not brand new. They did not compromise the ability to be wrung and measured.

The second participant, MIKES, reported slight scratching on steel GBs. However, they reported a slightly chipped edge on the 80 mm Ceramic GB that fortunately did not impede its measurement.

INRIM reported rust dots on the 0.5 mm, 25 mm, 80 mm and 100 mm blocks, and slight scratches on steel GBs. For the ceramic GBs they indicated that the 7 mm, 80 mm, and 90 mm had areas with missing material.

PTB indicated slight scratching of the steel GBs and indicated that the 80 mm ceramic GB was missing a small piece of material along the edge.

INMETRO acknowledged a small accident with the 7 mm steel GB that left a small burr. However, this did not alter the wringing capability or its measurement results.

NMISA did not report results on the 0.5 mm steel GB and did not comment on the reason for not measuring nor the quality of the block. NMIJ-AIST reported that they found the 0.5 mm GB to be in bad condition and that they could not wring it to a platen and measure it. However, the pilot laboratory was able to wring it and measure it at the end of circulation.

[11] The difficulty in wringing may be due to the use of platens of a different material, Tungsten Carbide.

The rest of the participants all reported slight scratches on the steel GB, and the missing a piece of material on the 80 mm Ceramic GB.

4.2 History of the gauge blocks

The length stability was to be assessed by the pilot laboratories by measuring three times: at the start, middle, and end of the Key Comparison exercise. Additionally, CENAM had a few previous available records of the short GBs used in the comparison. **Figures 4**, through **9** show all these measurements for the three longer short Steel GBs and the three longer ceramic GB that may be subject to noticeable drift. By observing these graphs, no clear drift may be identified. It must be mentioned at this point that the 2011 and 2012 pilot laboratory (CENAM) measurements might be slightly biased, as we experienced problems with our interferometer that were identified and corrected afterwards[12].

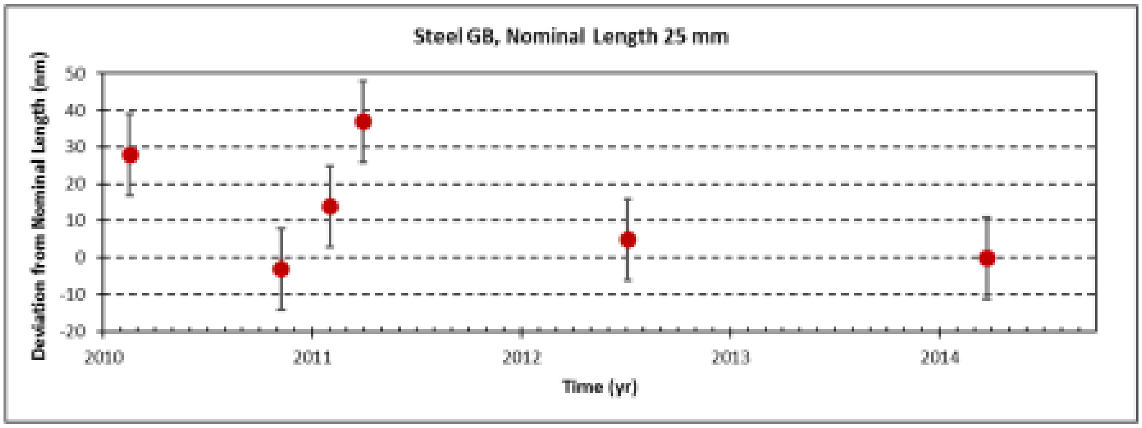

Figure 4. Historic measurements of the 25 mm Steel GB. In addition to these values, we have the manufacturer's Certificate of Inspection Value from a mechanical comparison measurement, (40 ± 37) nm, $k = 1$.

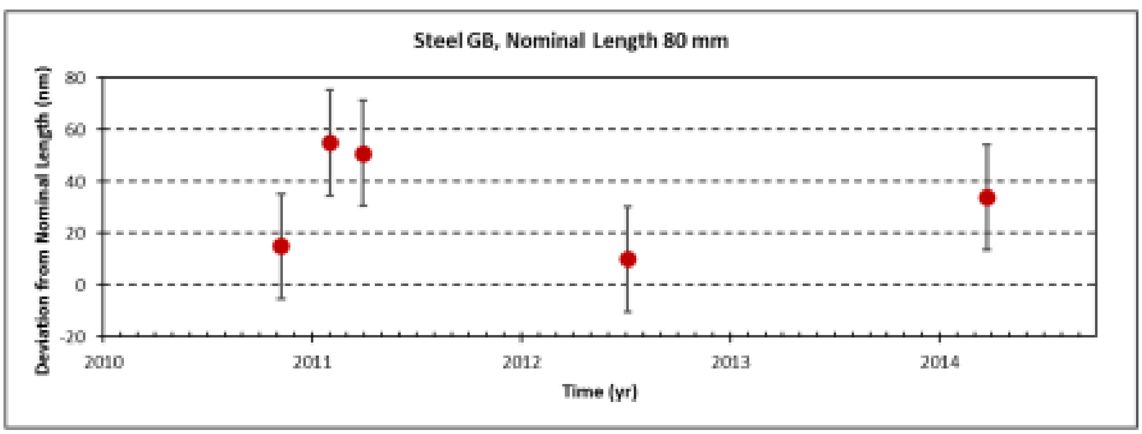

Figure 5. Historic measurements of the 80 mm Steel GB. In addition to these values, we have the manufacturer's Certificate of Inspection Value from a mechanical comparison measurement, (50 ± 50) nm, $k = 1$.

[12] A full explanation is given in Appendix A.

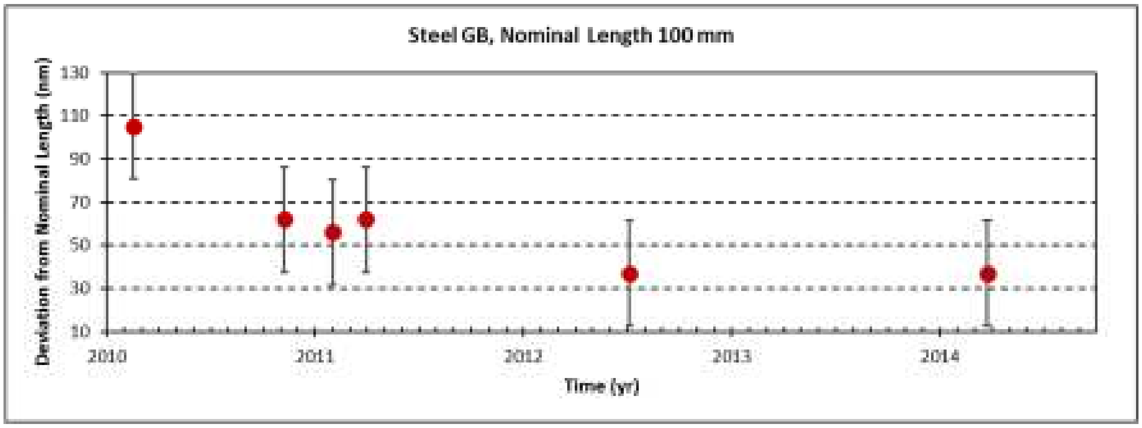

Figure 6. Historic measurements of the100 mm Steel GB. In addition to these values, we have the manufacturer's Certificate of Inspection Value from a mechanical comparison measurement, (100 ± 65) nm, $k = 1$.

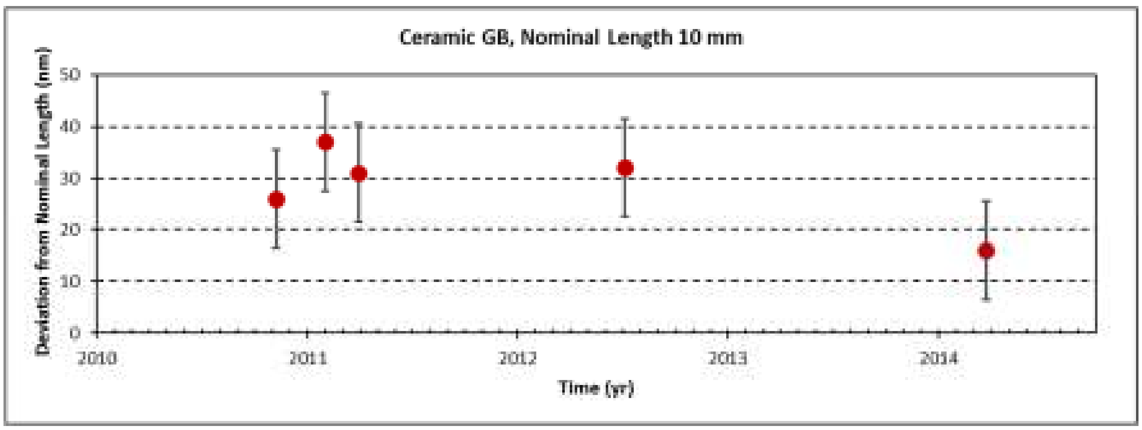

Figure 7. Historic measurements of the 10 mm Ceramic GB. In addition to these values, we have the manufacturer's Certificate of Inspection Value from a mechanical comparison measurement, (30 ± 37) nm, $k = 1$.

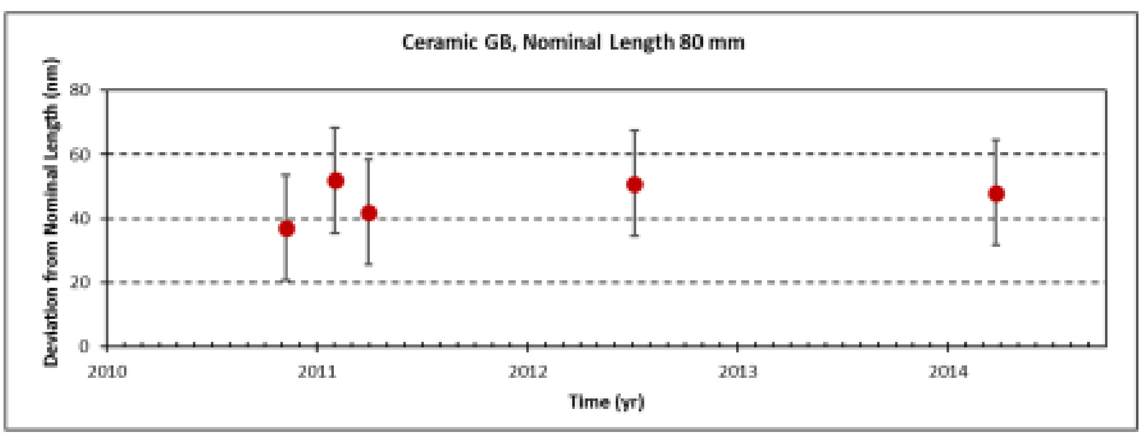

Figure 8. Historic measurements of the 80 mm Ceramic GB. In addition to these values, we have the manufacturer's Certificate of Inspection Value from a mechanical comparison measurement, (50 ± 50) nm, $k = 1$.

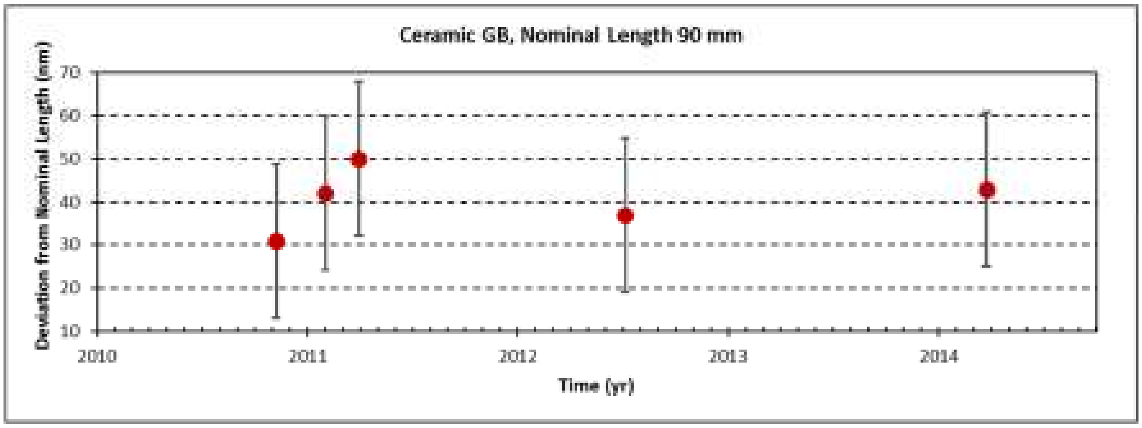

Figure 9. Historic measurements of the 90 mm Ceramic GB. In addition to these values, we have the manufacturer's Certificate of Inspection Value from a mechanical comparison measurement, (50 ± 65) nm, $k = 1$.

For the three long steel GBs, there were no historic records although they were not brand new and an attainment of a certain stability was expected. However, a discussion on drift of these GBs will be presented in Section 8.2.

Additionally, a temperature data logger was included with the short GBs to monitor the temperature during the whole circulation process. Unfortunately, the instrument stopped recording at the beginning of the second loop (probably somebody pressed the stop button on it) therefore there are recordings only of the first loop. **Figure 10** shows the temperature recording. However, no permanent change in the lengths of the artifacts at 20.0 C is expected from the GBs being exposed to these temperature variations.

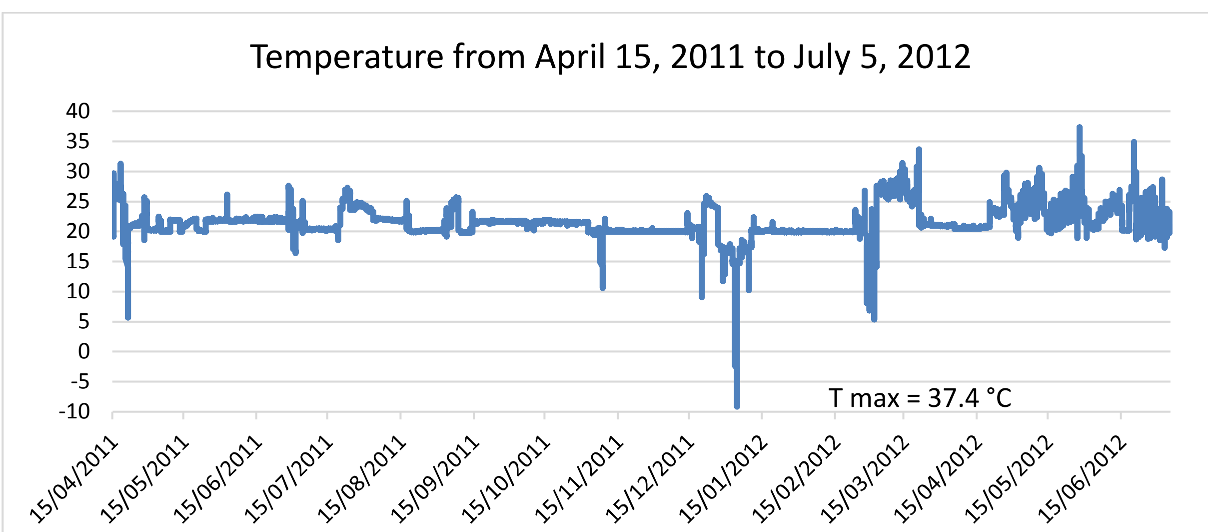

Figure 10. Data logger recording of temperature during circulation of the GB (only the first loop was recorded).

5. Definition of the measurands, calibration conditions, and uncertainty of measurements

5.1 Definition of the measurands

Calibration of short gauge blocks (0 mm to 100 mm) by optical interferometry requires the measurements to be performed in the vertical position, with the GBs wrung to a platen as indicated in Ref. [1]. The gauge block central length, l_c, is the measurand and is defined as the perpendicular distance between the central point of the free measurement surface of the gauge block and the surface of the platen.

The GBs should be measured wrung to the platens that the laboratory currently uses in their gauge block calibration services.

The length of long gauge blocks (over 100 mm) is defined in the horizontal position, with the blocks supported asymmetrically at two points to balance the weight of the platen wrung to one end [1]. If measured in this position the weight of the platen wrung to one end should be compensated. Otherwise, if they are measured vertically the deformation due to compression should be corrected.

The participants were instructed to measure both measuring faces of the GBs. The measurands that had to be reported were the deviations from nominal length (l_n) determined at the center for both measuring faces "A" and "B", $e_{cX} = l_c - l_n$, (where X = "A" or "B"); the average of both values, e_{avg}; the phase change correction, Δl_ϕ, that is described in Section 5.3; and the corrected average deviation after having applied the phase change correction, e_c. The measurement results were to be reported in the form provided in **Annex C1** for short GBs and in **Annex C2** for long GBs of the measurement protocol.

5.2 Reference temperature, standard pressure, and CTE

The measurement results were to be reported at the reference temperature of 20 °C and standard pressure of 101 325 Pa. The CTE's provided in this document were to be used to make corrections. Additional corrections may be applied according to the specific procedure of each laboratory.

5.3 Optical phase change correction on the reflection of light

The position of the plane where light is reflected on a surface differs from the position of mechanical contact of the surface. This difference varies depending mainly on the material and surface finish of the surface. As the free measuring face of the GB under measurement and the platen where it is wrung are in general different in both characteristics, this difference varies and a correction has to be applied. It is called the phase change correction, Δl_ϕ. This correction, if considered, had to be reported using **Annex C** of the protocol as an individual value for each short GB material, and, eventually, for the long GB. If reported, it had to be estimated or determined by each laboratory according to its calibration procedure as is usually done for its customers. The most common method to determine the phase change correction is the stack method described in the measurement protocol. It was used by all participants with the exception of NIST and PTB. The material of the platens used to wring the GBs and the values obtained by the participant NMI are listed in **Table 7**.

The laboratories used different kinds of GB interferometers. **Table 8** shows the description they provided of their systems.

	Steel gauge blocks		**Ceramic gauge blocks**	
	Platen Material	**Value (nm)**	**Platen Material**	**Value (nm)**
CENAM	steel	-35 (-20 for 0.5 mm)	steel	-23
NRC-CNRC	fused silica steel	+32 (+31 0.5 mm) + 18 nm	fused silica	+50
METAS	tungsten carbide	+2	tungsten carbide	+18
MIKES	steel	+6.6	steel	+11
INRIM	ceramics	+3	ceramics	+4
PTB	crystalline quartz, steel	Individual GB values[13]: Nom. Val. (mm) – (nm) 0.5 – 27.5 3 – 0.5 5 – 1.0 7 – 0.5 25 – -1.5 80 – -1.0 100 – 0.5	ceramics	0[14]
NIST	steel	+5.4[15]	steel	-7.0[15]
INMETRO	quartz	-+31 (for the 25, 80 and 100 mm) -+38 (GB 0.5, 3, 5 and 7 mm)	quartz	+48
NMIJ-AIST	steel	0	ceramics	0
NIM	glass	+59	glass	+32
A*STAR	steel	-15.2	steel	-29.8
NMISA	quartz	Did not inform	quartz	Did not inform

Table 7. Phase Change Correction applied by the NMIs to the two GB materials.

[13] Individual measurements by scattered light technique with integrating sphere (roughness contribution). In case of dissimilar materials of GB and platen (quartz), the phase shift on reflection for the GB is considered (data pools of previous investigations).

[14] Individual measurements are not intended in routine PTB calibration procedures for GBs and platens made of ceramics. Phase Change Corrections were not taken in account, although significant phase change/roughness differences between GBs and platens made of ceramics may exist. Instead, the length independent measurement uncertainty contribution was significantly increased compared with the steel GBs to take in account the lack of correction. Standard uncertainty from 14 to 15 nm for ceramic GB while 9.5 to 10.5 nm for steel GB.

[15] Spherical contact technique.

5.4 Mathematical model and measurement uncertainty

All participating laboratories should have a mathematical model they apply to describe their measuring process corresponding to their measuring procedure and from which their uncertainty budget is established, according to [2].

A general mathematical model for the central length deviation, e_c, proposed in [3] was enunciated in the protocol. However, depending on the specific method, equipment and considerations used by each laboratory it could be different.

The laboratories had to apply the mathematical model they use according to their measurement method along with the list of uncertainty sources considered in their procedures. They were requested to submit the mathematical model equation. According to the submitted model, they were required to fill out an uncertainty budget chart. Both, the mathematical model as well as the uncertainty budget chart were requested in the **Annex D** form.

6. Measurement results and declared uncertainties.

6.1 Measurement results

The measurement results of e_c, central length of each gauge block, for the participating NMIs are summarized in the following three tables: **Table 9** for the short steel gauge blocks, **Table 10** for the short ceramic gauge blocks, and **Table 11** for the long steel gauge blocks. Participants appear in the order of participation.

Table 8.	Make and type of GB interfer.	Light sources, Wave lengths	Method of fringe fraction determination	Refract. index det.	Temperature range (°C)
CENAM	NPL TESA, Twyman Green	HeNe Laser TESA, 633 nm HeNe Laser TESA, 543 nm	Fringe Fractions evaluated by software from a digitalized image. Both wavelengths are taken into account for wave number determination and fringe fraction determination.	Edlén	From 19.23 to 20.12 variation.
NRC-CNRC	Home Made (NRC-CNRC), Twyman Green Interferometer	HeNe Laser Coherent, 633 nm HeNe Laser TESA, 612 nm HeNe Laser TESA, 543 nm Home-made polarization-satbilized laser 1152 nm	Measurement of fringe fractions using the image produced by a video camera and by visual interpolation with a fiducial	Ciddor´s equation	The maximum absolute temperature variations from nominal were, on average, 0.011 K with an uncertainty on the mean of ± 0.002 K. The largest single measurement deviation from 20 °C was 0.017 K with a standard uncertainty of ± 0.006 K.
METAS	NPL TESA gauge block Interferometer	Red HeNe stabilized laser ,λ = 633 nm Green HeNe Zeeman stabilized laser ,λ = 543 nm	Video fringe analysis	Edlén	Temperature variation during measurements ± 0.05 °C at most
MIKES	NPL TESA gauge block Interferometer	Red 633 nm HeNe stabilized laser Green 543.5 nm HeNe stabilized laser	From digitized still images with suitable fringe pattern for both red and green illumination	Edlén	From 19.953 to 20.063 variation.
INRIM	Modified and improved Fizeau interferometer by Hilger & Watts	Stabilized He-Ne Laser 633 nm ^{86}Kr lamp: red 646 nm, orange 606 nm and violet 450 nm	The CCD images of interferograms are processed by software: (i) each fringe is substituted by a representative line; (ii) the fractional shift of the front to the rear interferogram is computed at the measuring face centre.	Edlén	From 19.90 to 20.04 variation.
PTB	Michelson-interferometer, modified TSUGAMI (Japan), supplied with phase stepping arrangement, automatic evaluation and active temperature control.	780 nm Rd-stabilized diode-laser 633 nm frequency stabilized HeNe laser 532 nm iodine stabilized Nd:YAG laser	Automatic interference evaluation by phase stepping interferometry, computer-aided length determination, topography.	Edlén	From 19.987 to 20.016 variation for steel GB From 19.987 to 20.010 variation for ceramic GB
NIST	NPL Hilger Gauge Interferometer No. 1955 for short GB; Moore CMM coupled to a heterodyne interferometer for long GB.	Spectra Physics 117A stabilized HeNe Red; 633 nm heterodyne interferometer.	Visual interpolation of two operators assisted by a Gaertner Etalon Scale for short GB; for long GB, the GB is wrung to a platen and the the platen and the GB measuring face are probed with the CMM Meas. Head. The displacement is measured with a heterodyne interferometer.	Edlén	From 20.050 to 20.100 variation.
INMETRO	Jena-Zeiss vertical, Twyman-Green type (25 mm to 100 mm GB) GBI-Mitutoyo automated vertical, Twyman-Green type (0.5 mm to 25 mm GB)	^{114}Cd spectral lamp (red, green, blue, violet) HeNe stabilized lasers (green and red)	Fringe fraction automated (2D) estimation in 4 wavelengths (red, green, blue, violet) Fringe fraction automated (phase –stepping) estimation in 2 wavelengths (red, green)	Birch-Downs air refractive index calculation	Less than 0.02 °C variation.in block/air
NMIJ-AIST	NRLM-TSUGAMI, Twyman-Green interferometer	Zeeman stab. laser, 633 nm (for GB ≤ 152.4 mm); I_2 stab. offset-locked laser, 633 nm (for GB > 152.4 mm); I_2 stab. SHG of Nd:YAG, 532 nm; and Rb stab. laser diode, 780 nm.	By image processing of an interference image. The interference fringes were generated by tilting a gauge block and platen. The fringes on the GB and the platen were fitted to sinusoidal functions and their phase difference was calculated.	Ciddor´s equation	From 19.9 to 20.1 variation.
NIM	An improved Koester interferometer made by Zeiss-Jena, the former East Germany	633 nm laser	Phase-stepping fringe fractions measurement at one wavelength	Edlén	From 19.9 to 20.1 variation.
A*STAR	NPL-TESA Automatic Gauge Block Interferometer	Stabilised He-Ne Red Laser Stabilised He-Ne Green Laser	The resulting interference fringe pattern was captured by the CCD camera in the system. The fringe fraction number was determined by analysing the intensity profile of the displaced fringes.	Edlén	From 19.8 to 20.3 variation.
NMISA	TESA GBI 300 gauge block interferometer	Red and Green wave lengths	According to TESA procedure, Tesa manufacturer software	TESA system	From 19.6 to 20.3 variation.

Steel Short Gauge Blocks							
Nominal Length (mm)	**0.5**	**3**	**5**	**7**	**25**	**80**	**100**
NMI	**Deviation from Nominal Central Length, e_c (nm)**						
CENAM	-17	-22	40	27	37	51	62
NRC-CNRC	-26	-43	5	-4	17	33	35
METAS	-32	-36	36	-4	-12	24	37
MIKES	-13	-31	11	4	17	23	41
INRIM	-20	-28	7	11	7	11	29
PTB	-31	-42	-5	-9	-1	20	40
NIST	0.4	-7.6	28.4	16.4	18.4	28.4	55.4
INMETRO	-32	-20	14	-5	12	47	44
CENAM	*-26*	*-38*	*35*	*10*	*5*	*10*	*37*
NIM	-26	-34	18	12	21	34	47
A*STAR	-16.2	-27.7	11.8	39.3	13.8	15.3	30.8
NMISA	-[16]	-10	6	-2	8	22	25
NMIJ/AIST	-[15]	-37	5	-2	-2	11	30
CENAM	*-20*	*-32*	*2*	*1*	*0*	*34*	*37*

Table 9. Participants central length deviation from nominal value, e_c, measurement results for the seven short steel gauge blocks. Participants appear in the order of participation. Values in italics are the control measurements made by the pilot laboratory.

Ceramic Short Gauge Blocks						
Nominal Length (mm)	**3**	**5**	**7**	**10**	**80**	**90**
NMI	**Deviation from Nominal Central Length, e_c (nm)**					
CENAM	3	43	60	31	42	50
NRC-CNRC	-5	51	63	37	70	64
METAS	-16	56	60	33	68	59
MIKES	-12	34	54	31	52	53
INRIM	-21	37	43	21	48	30
PTB	-27	33	35	10	57	45
NIST	-13	39	45	21	43	50
INMETRO	-5	54	29	43	69	78
CENAM	*-5*	*45*	*51*	*32*	*51*	*37*
NIM	-30	23	39	18	29	21
A*STAR	-21.8	29.2	41.2	30.7	0.2	-7.8
NMISA	-42	11	13	6	72	67
NMIJ/AIST	-1	53	59	36	67	56
CENAM	*-29*	*34*	*30*	*16*	*48*	*43*

Table 10. Participants central length deviation from nominal value, e_c, measurement results for the six short ceramic gauge blocks. Participants appear in the order of participation. Values in italics are the control measurements made by the pilot laboratory.

[16] This participant did not report results on this gauge block.

Steel Long Gauge Blocks			
Nominal Length (mm)	152.4	254	508
NMI[17]	**Dev. from Nom. Cent. Length, e_c (nm)**		
NRC-CNRC	246	743	1772
METAS	194	734	1794
MIKES	171	719	1773
INRIM	210	770	1810
PTB	185	738	1820
NIST	189	700	1786
INMETRO	221	781	1868
NRC-CNRC	*207*	*769*	*1857*
NIM	174	722	1792
NMISA	240	770	1745[18]
NMIJ/AIST	210	816	1925
CENAM	207	770	-[19]
NRC-CNRC	*221*	*781*	*1881*

Table 11. Participants central length deviation from nominal value, e_c, measurement results for the three long steel gauge blocks. Participants appear in the order of participation. Values in italics are the control measurements made by the pilot laboratory.

6.2 Declared standard uncertainties.

The declared Combined Standard Uncertaities, *u*, for each gauge block are shown on **Tables 12**, **13**, and **14** for, respectively, the short steel gauge blocks, ceramic gauge blocks, and long steel gauge blocks.

Steel Short Gauge Blocks							
Nominal Length (mm)	**0.5**	**3**	**5**	**7**	**25**	**80**	**100**
NMI	**Declared Combined Standard Uncertainties, *u*(nm)**						
CENAM	9.3	9.4	9.4	9.4	10.4	17.2	20.3
NRC-CNRC	19	15	15	15	15	18	17
METAS	10	10	18	10	10	13	16
MIKES	10	10	10	10	11	15	17
INRIM	9	9	9	9	10	16	19
PTB	9.5	8.5	8.5	8.5	8.5	9.5	10.5
NIST	9	9.2	9.4	9.5	11.2	16.1	17.9
INMETRO	12	12	12	12	13	16	17
CENAM	*9.3*	*9.4*	*9.4*	*9.4*	*10.4*	*17.2*	*20.3*
NIM	7	7	7	7	7	10	12
A*STAR	14.4	14.4	14.5	14.6	16.6	30	35.9
NMISA	-	13	13	13	15	26	31
NMIJ	-	13	13	19	13	21	22
CENAM	*9.3*	*9.4*	*9.4*	*9.4*	*10.4*	*17.2*	*20.3*

Table 12. Declared combined standard uncertainties, *u*, for the seven short steel gauge blocks.

[17] A*STAR is excluded because they do not offer the service of long gauge block calibration.
[18] 1745 nm obtained from average value corrected for phase change. However, 1710 nm was reported.
[19] This laboratory only offers gauge block calibration up to 300 mm.

Ceramic Short Gauge Blocks						
Nominal Length (mm)	**3**	**5**	**7**	**10**	**80**	**90**
NMI	**Declared Combined Standard Uncertainties, *u*(nm)**					
CENAM	9.4	9.4	9.4	9.5	16.5	17.9
NRC-CNRC	14	15	15	15	15	15
METAS	13	13	13	13	15	16
MIKES	10	10	10	10	15	16
INRIM	9	9	9	9	16	18
PTB	14	14	14	14	15	15
NIST	9.2	9.4	9.5	9.8	16.1	17
INMETRO	12	12	12	12	15	16
CENAM	*9.4*	*9.4*	*9.4*	*9.5*	*16.5*	*17.9*
NIM	7	7	7	7	10	11
A*STAR	14.4	14.5	14.6	14.7	27.9	30.5
NMISA	13	13	13	13	26	28
NMIJ	13	13	13	13	15	15
CENAM	*9.4*	*9.4*	*9.4*	*9.5*	*17.2*	*18.7*

Table 13. Declared combined standard uncertainties, *u*, for the six short ceramic gauge blocks.

Steel Long Gauge Blocks			
Nominal Length (mm)	152.4	254	508
NMI	**Declared Combined Standard Uncertainties, *u* (nm)**		
NRC-CNRC	20	22	31
METAS	23	21	31
MIKES	17	20	32
INRIM	29	31	39
PTB	14	14.5	20.5
NIST	35.9	42.7	59.75
INMETRO	10	13	24
NRC-CNRC[20]	*18*	*20*	*33*
NIM	17	25	46
NMISA	44	71[21]	320[22]
NMIJ	21	26	44
CENAM	30	48	-
NRC-CNRC[19]	*18*	*24*	*40*

Table 14. Declared combined standard uncertainties, *u*, for the three steel long gauge blocks

[20] Did not measure again the Phase Change Correction.
[21] This value exceeds the limit required uncertainty for this length, 61 nm.
[22] This value exceeds the limit required uncertainty for this length, 102 nm.

7. Determination of the reference values

7.1 Weighted mean as the key comparison reference value

It was decided by the sWG-Key Comparisons to use the weighted mean as the key comparison reference value (KCRV) as it is a general practice for key comparisons. The weighted mean, $\bar{x}_W$, is given by equation (1) and its corresponding uncertainty by equation (2):

$$\bar{x}_W = \frac{\sum_{i=1}^{n} u^{-2}(x_i) \cdot x_i}{\sum_{i=1}^{n} u^{-2}(x_i)} , \tag{1}$$

$$u(\bar{x}_W) = \left(\sum_{i=1}^{n} u^{-2}(x_i) \right)^{-1/2} , \tag{2}$$

where $u(x_i)$ is the standard uncertainty of the measurement x_i.

7.2 Weighted mean values of the gauge blocks

In **Table 15** the weighted mean values obtained for the steel short gauge blocks, the ceramic gauge blocks and the steel long gauge blocks, along with their corresponding standard uncertainties are listed considering the contribution of all participants to the KCRV.

Nominal Length (mm)	Steel Short Gauge Blocks (nm)		Nominal Length (mm)	Ceramic Gauge Blocks (nm)		Nominal Length (mm)	Steel Long Gauge Blocks (nm)	
	$\bar{x}_W$	$u(\bar{x}_W)$		$\bar{x}_W$	$u(\bar{x}_W)$		$\bar{x}_W$	$u(\bar{x}_W)$
0.5	-20.9	3.1	**3**	-16.7	3.1	**152.4**	203.1	5.6
3	-28.6	2.9	**5**	36.5	3.1	**254**[23]	752.3	6.7
5	14.7	3.0	**7**	45.0	3.1	**508**[18]	1818.4	10.5
7	7.5	3.0	**10**	25.2	3.1			
25	11.7	3.1	**80**	51.2	4.4			
80	27.4	4.3	**90**	47.4	4.7			
100	41.5	4.9						

Table 15. Weighted mean with the corresponding standard uncertainty for the steel short gauge blocks, ceramic gauge blocks and steel long gauge blocks. All values in nm.

[23] NMISA was excluded from the calculation of the weighted mean because the declared uncertainty exceeded the limit required in section 2.1.

8 Measurement results and reference values

8.1 Graphs of measurement results and weighted mean

The results of participants along with the weighted mean considering all participants are shown in this section, from **Fig. 11** to **Fig. 26**. At this stage, no outliers' analysis was performed. The weighted mean was calculated considering all participants that submitted results for the corresponding gauge block. Additionally, no drift on any artifact is considered either. The participants' results are shown along with their corresponding standard uncertainty bars. The weighted mean is shown as a solid red line and its standard uncertainty limits are depicted by red dashed lines. The pilot laboratories results are their first measurements in all cases. The other two control measurements are shown but do not contribute to the weighted mean.

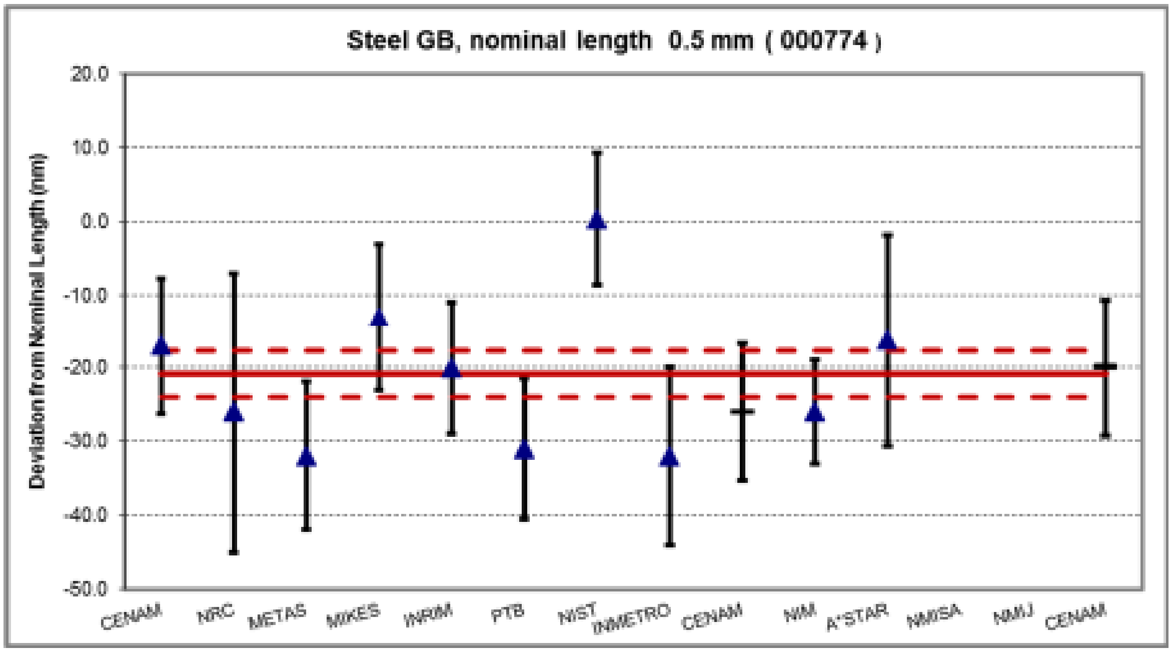

Figure 11 Measurement results of the participants along with the weighted mean for the 0.5 mm steel gauge block (GB). All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the upper and lower limits of its standard uncertainty. The laboratories NMISA and NMIJ did not report results for this GB. Values shown with a dash represent the second and third measurements by the pilot laboratory, CENAM, which were excluded from the calculation of the weighted mean.

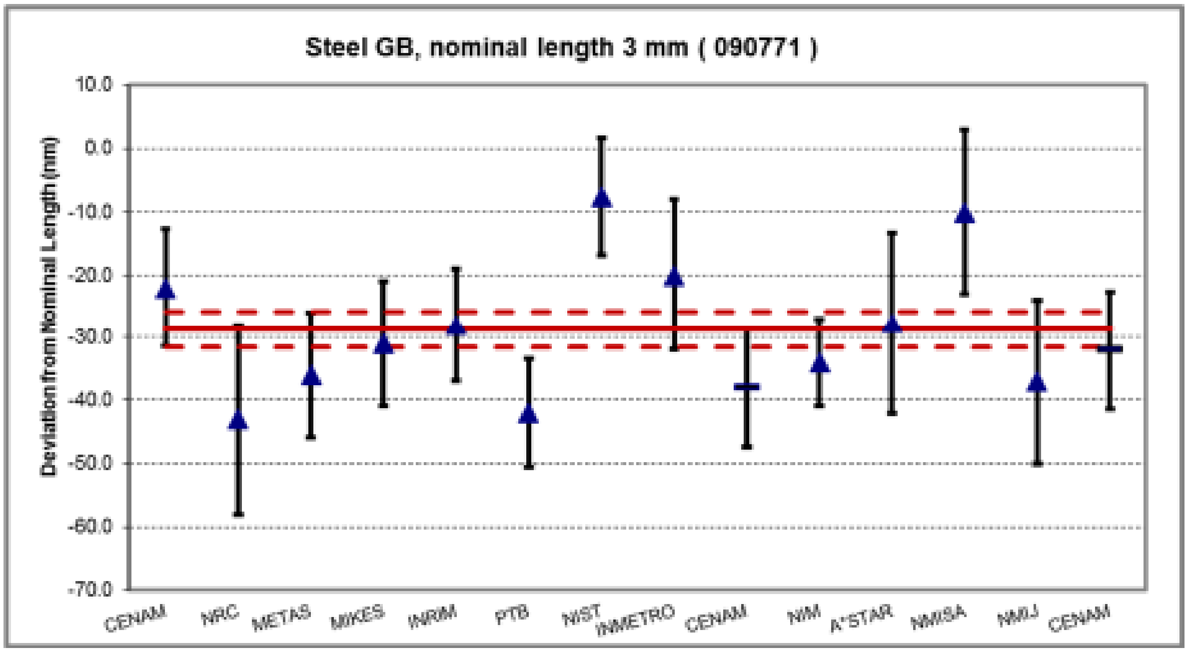

Figure 12 Measurement results of participants along with the weighted mean for the 3 mm steel GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the upper and lower limit of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

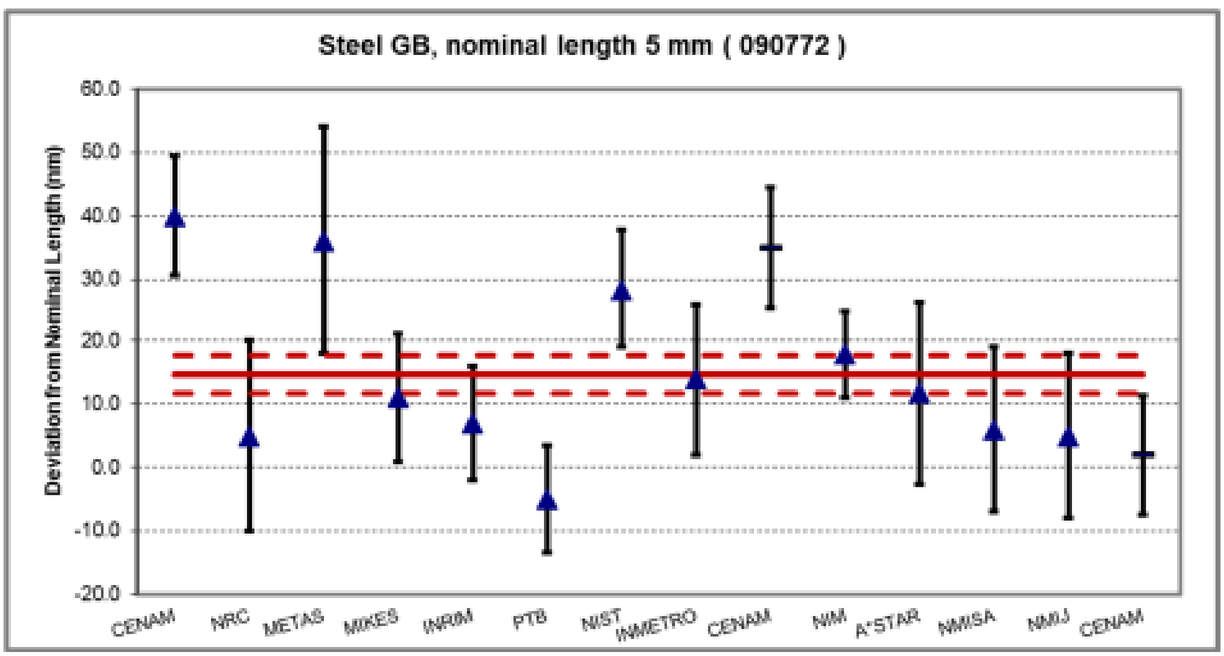

Figure 13 Measurement results of participants along with the weighted mean for the 5 mm steel GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the upper and lower limit of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

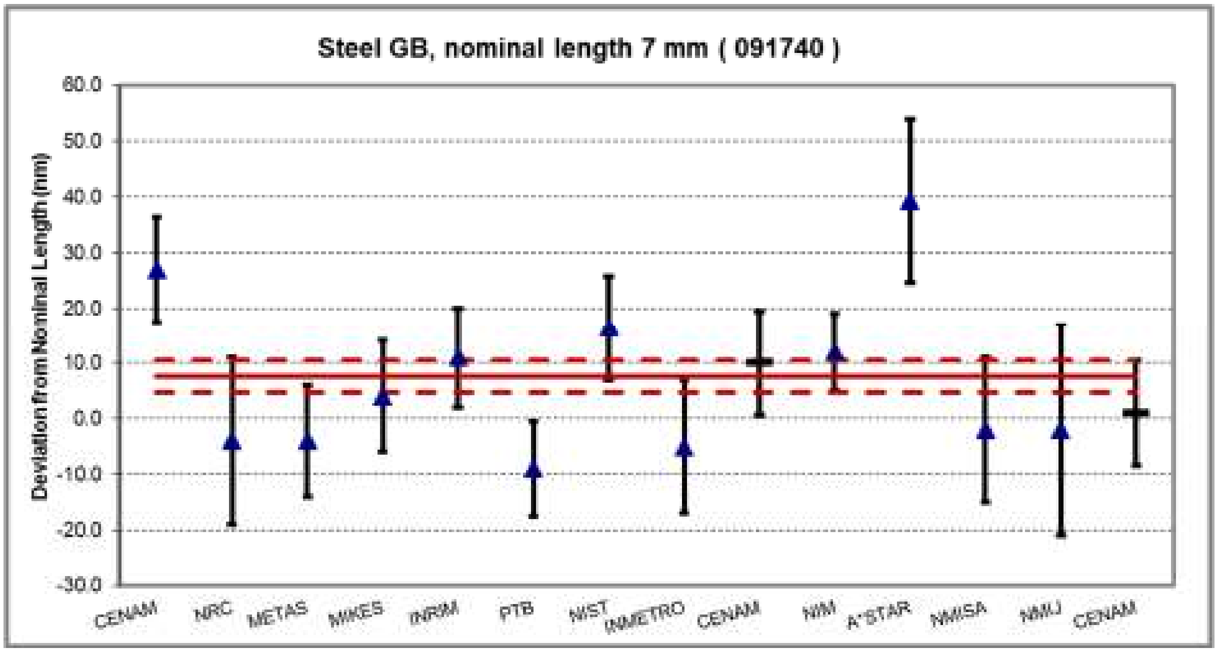

Figure 14 Measurement results of participants along with the weighted mean for the 7 mm steel GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

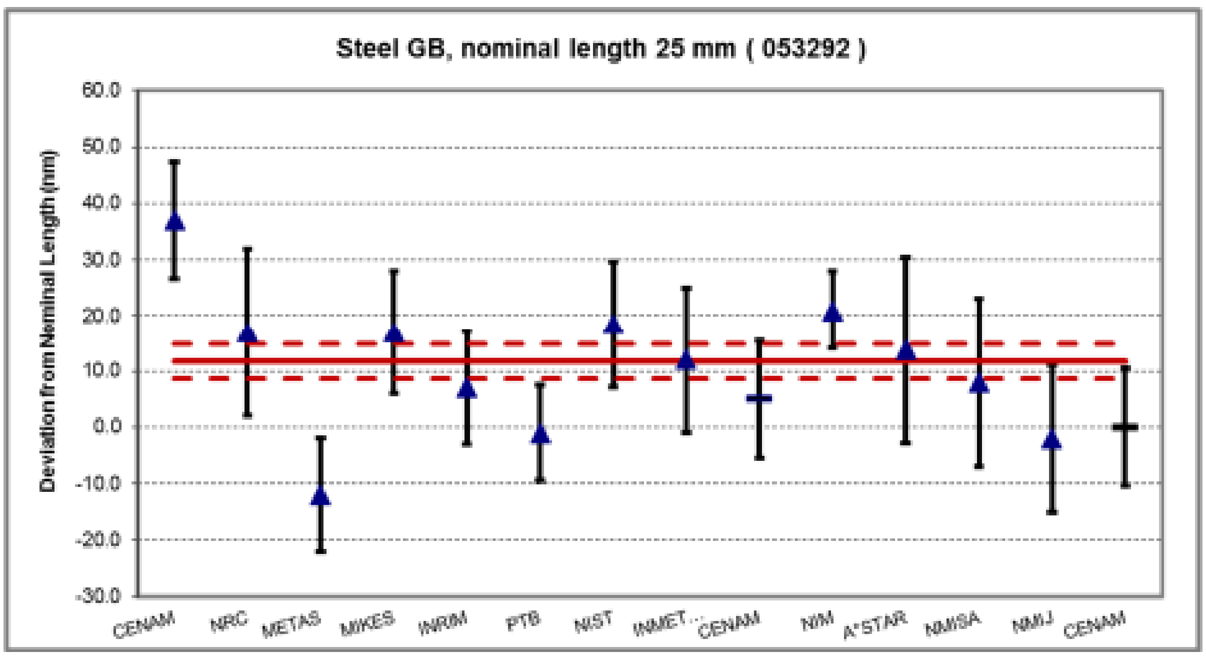

Figure 15 Measurement results of participants along with the weighted mean for the 25 mm steel GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the upper and lower limit of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

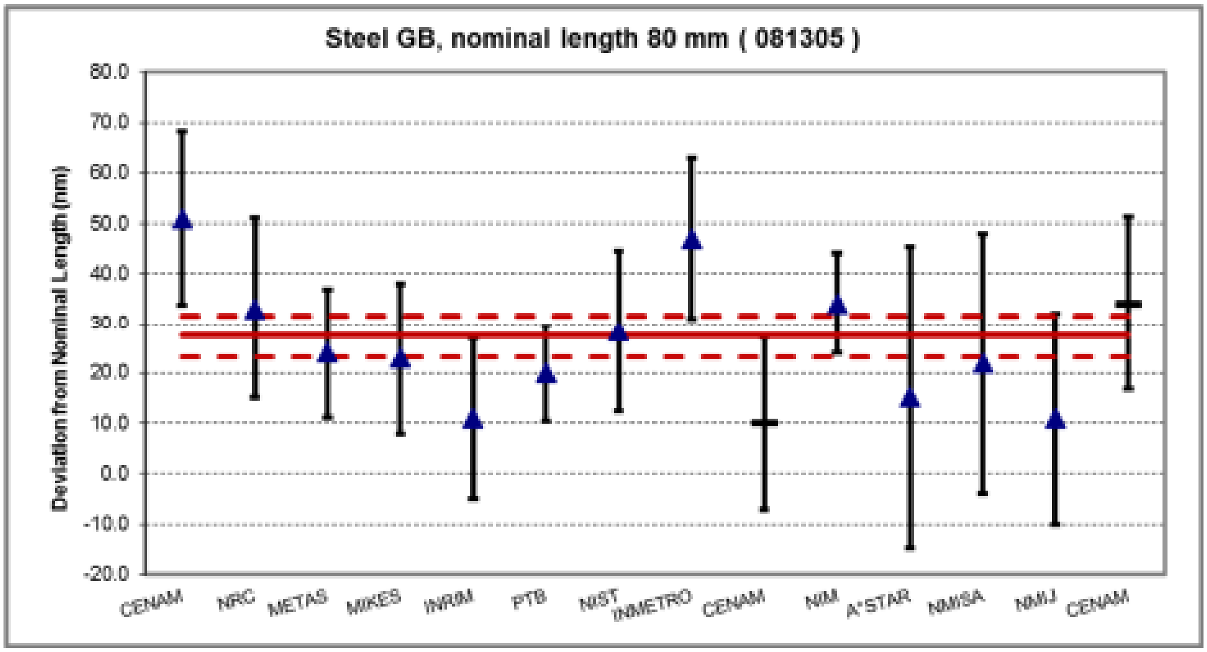

Figure 16 Measurement results of participants along with the weighted mean for the 80 mm steel GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

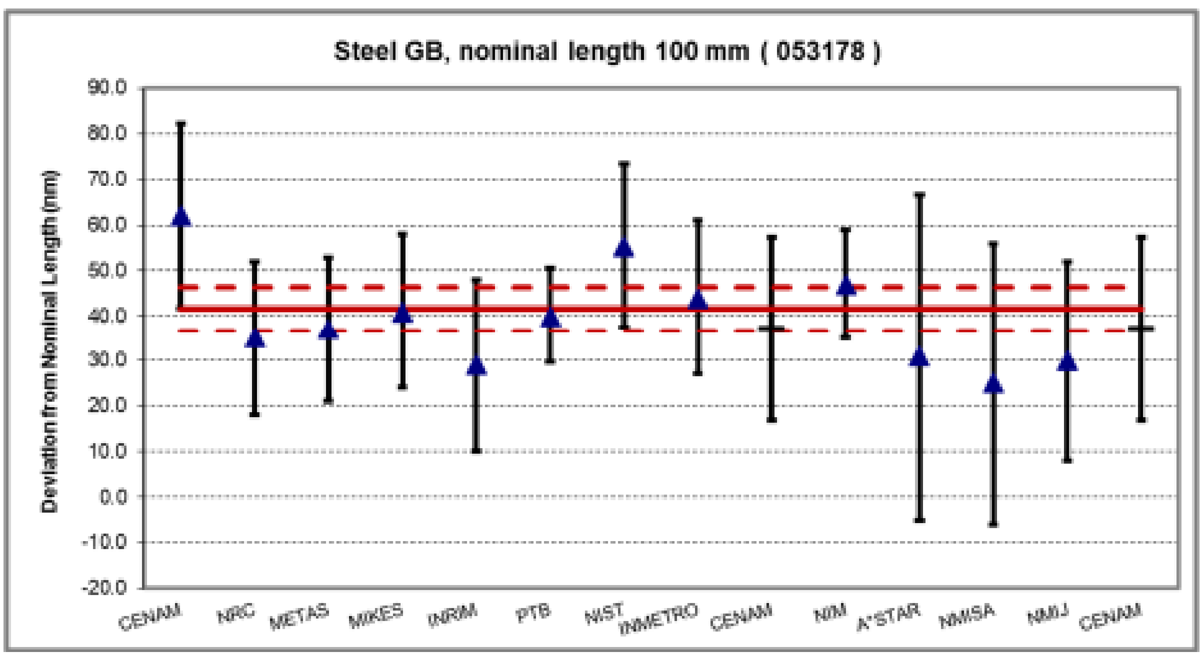

Figure 17 Measurement results of participants along with the weighted mean for the 100 mm steel GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

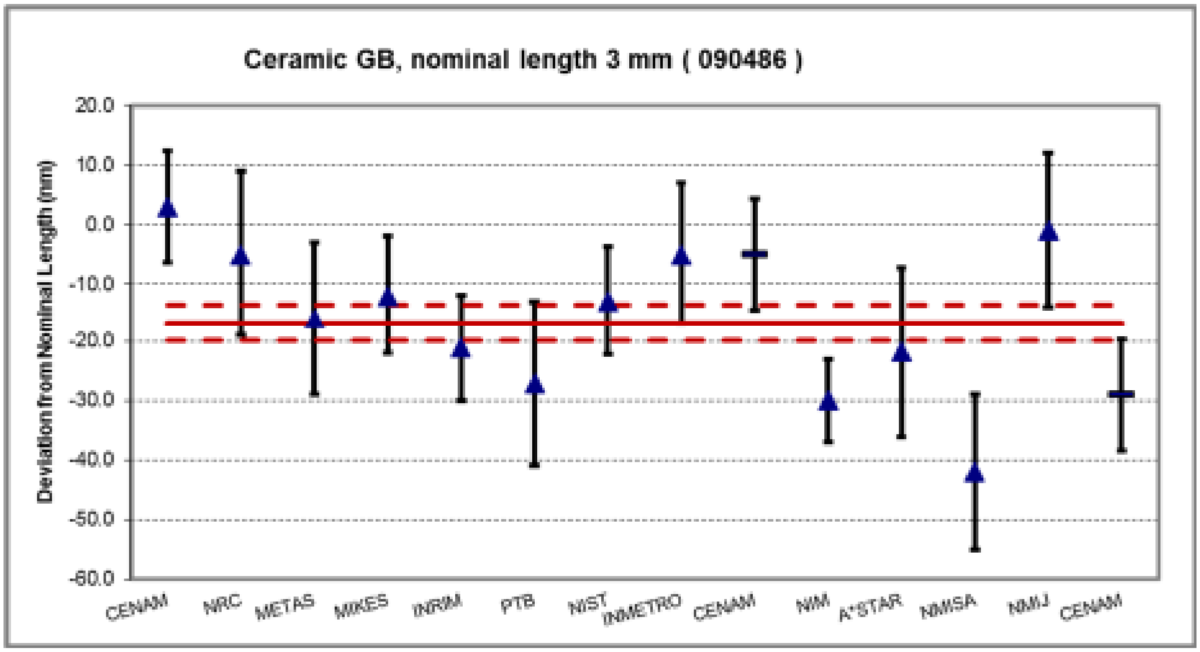

Figure 18 Measurement results of participants along with the weighted mean for the 3 mm ceramic GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory, CENAM, which were excluded from the calculation of the weighted mean.

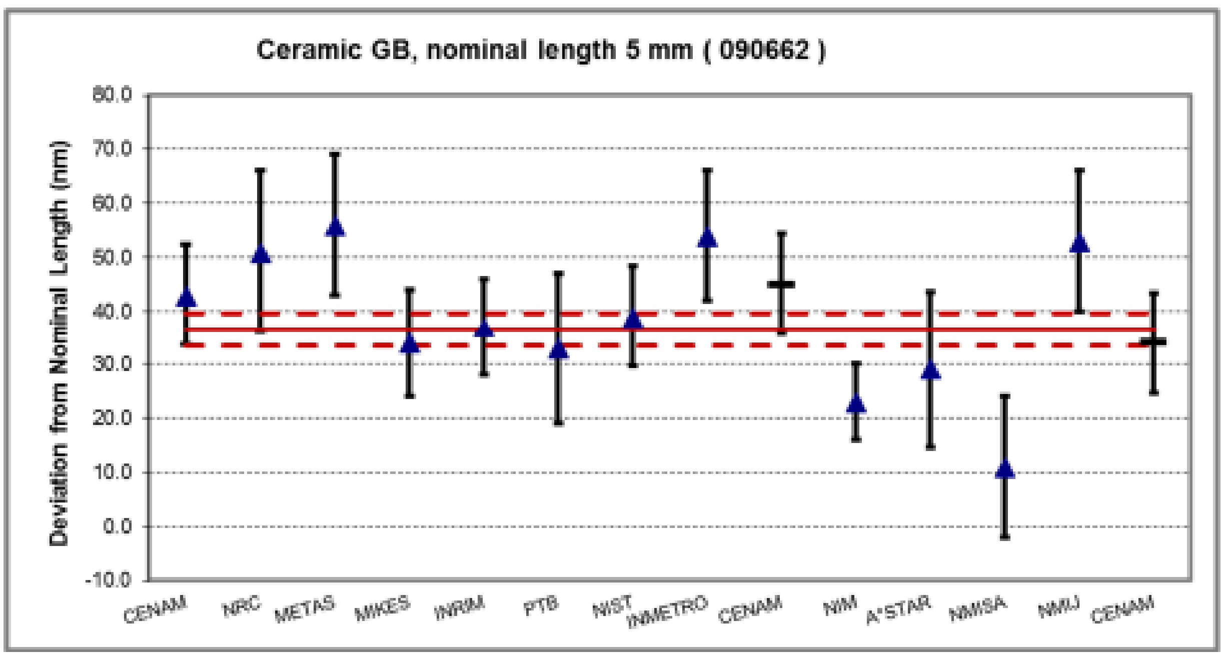

Figure 19 Measurement results of participants along with the weighted mean for the 5 mm ceramic GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

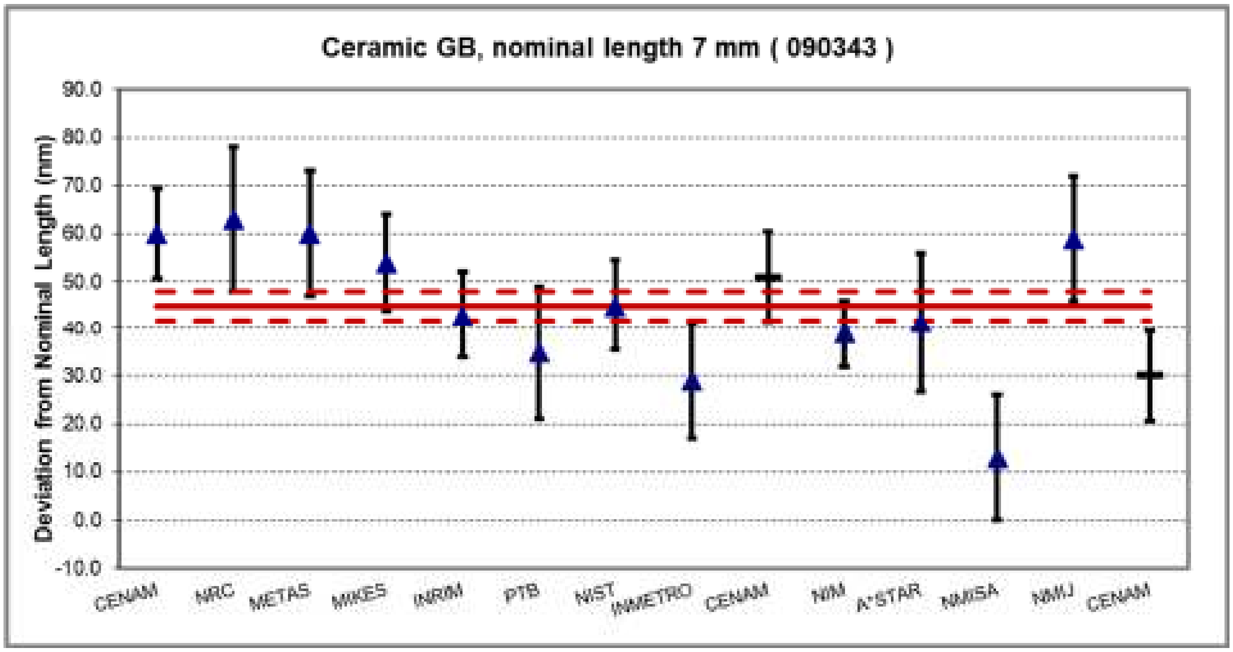

Figure 20 Measurement results of participants along with the weighted mean for the 7 mm ceramic GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

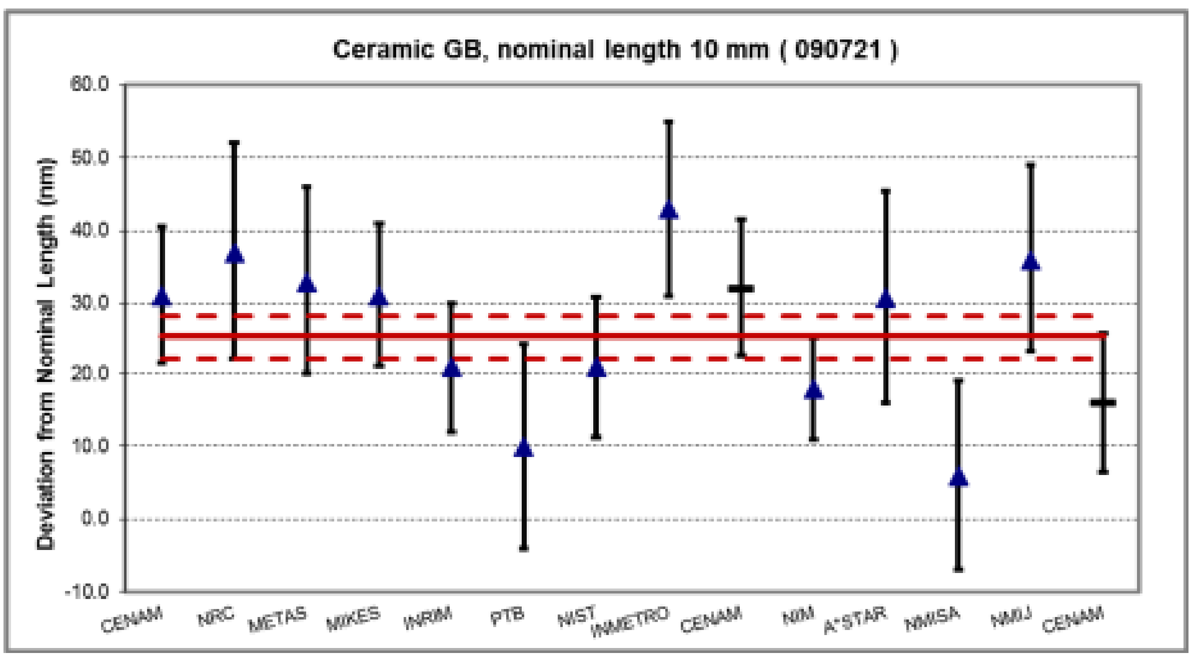

Figure 21 Measurement results of participants along with the weighted mean for the 10 mm ceramic GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

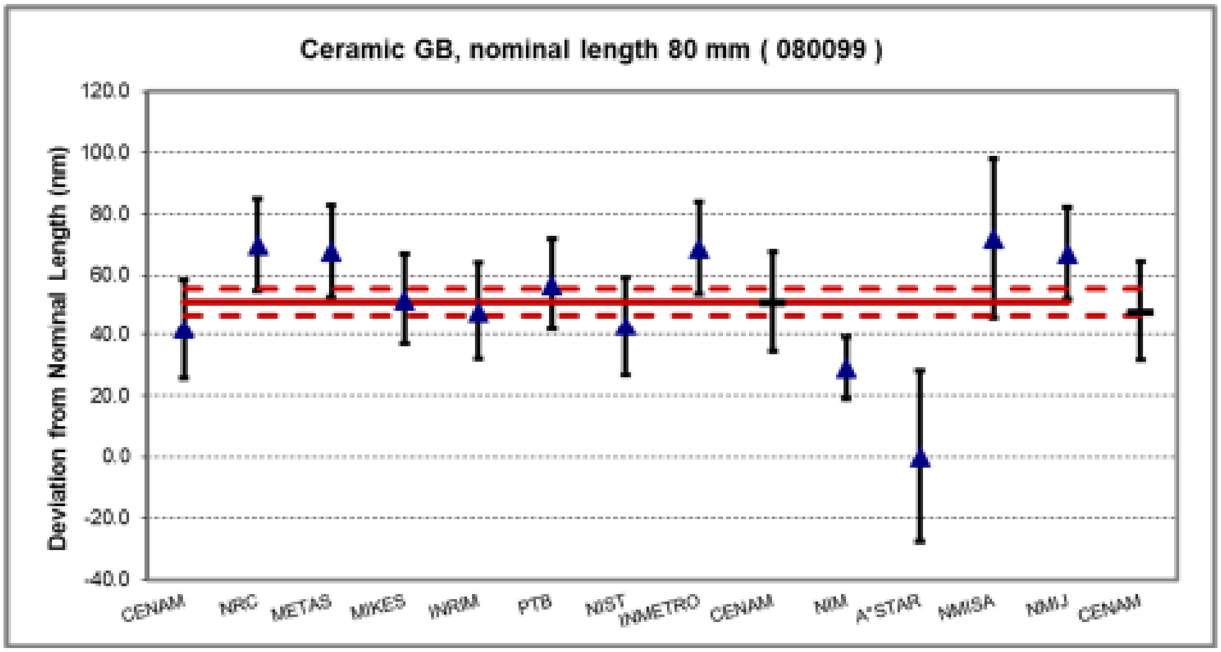

Figure 22 Measurement results of participants along with the weighted mean for the 80 mm ceramic GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

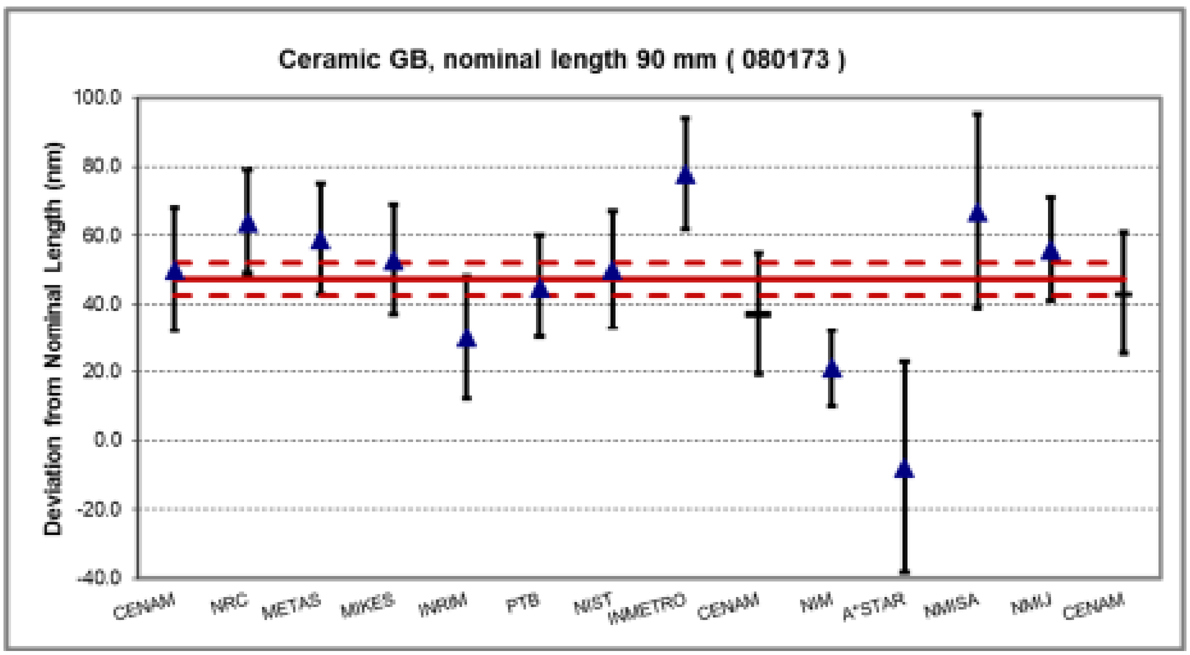

Figure 23 Measurement results of participants along with the weighted mean for the 90 mm ceramic GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines limits of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

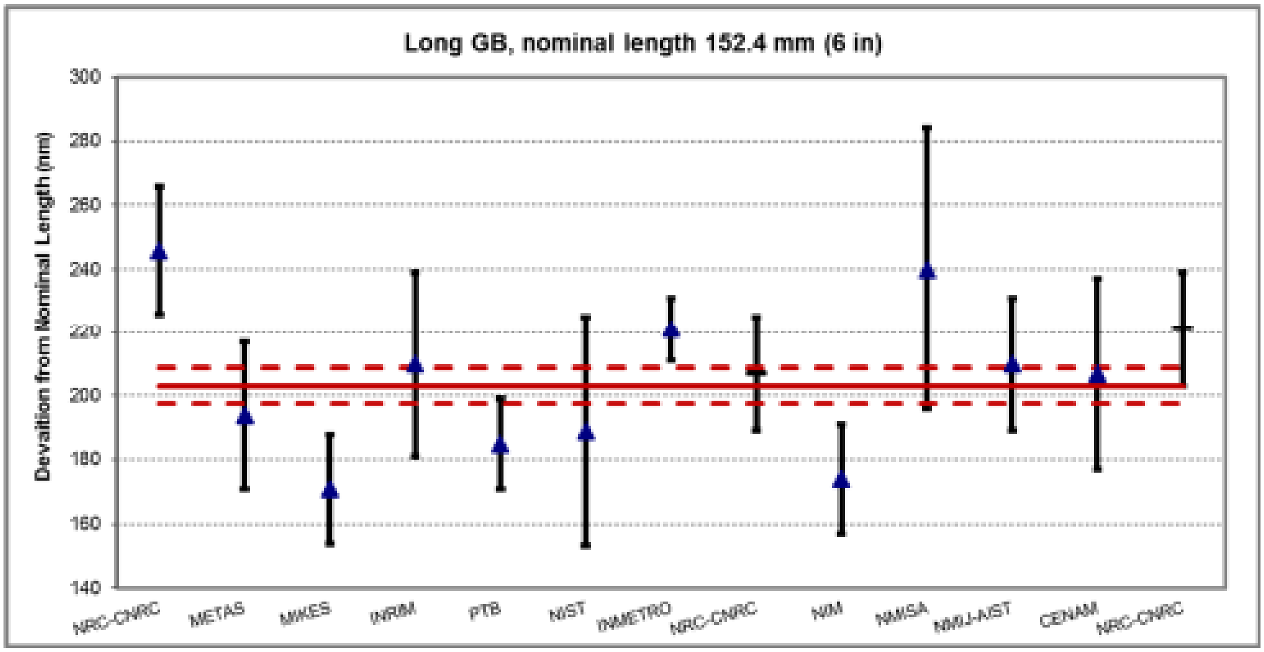

Figure 24 Measurement results of participants along with the weighted mean for the 152.4 mm long GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the upper and lower limit of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory, NRC-CNRC, which were excluded from the calculation of the weighted mean.

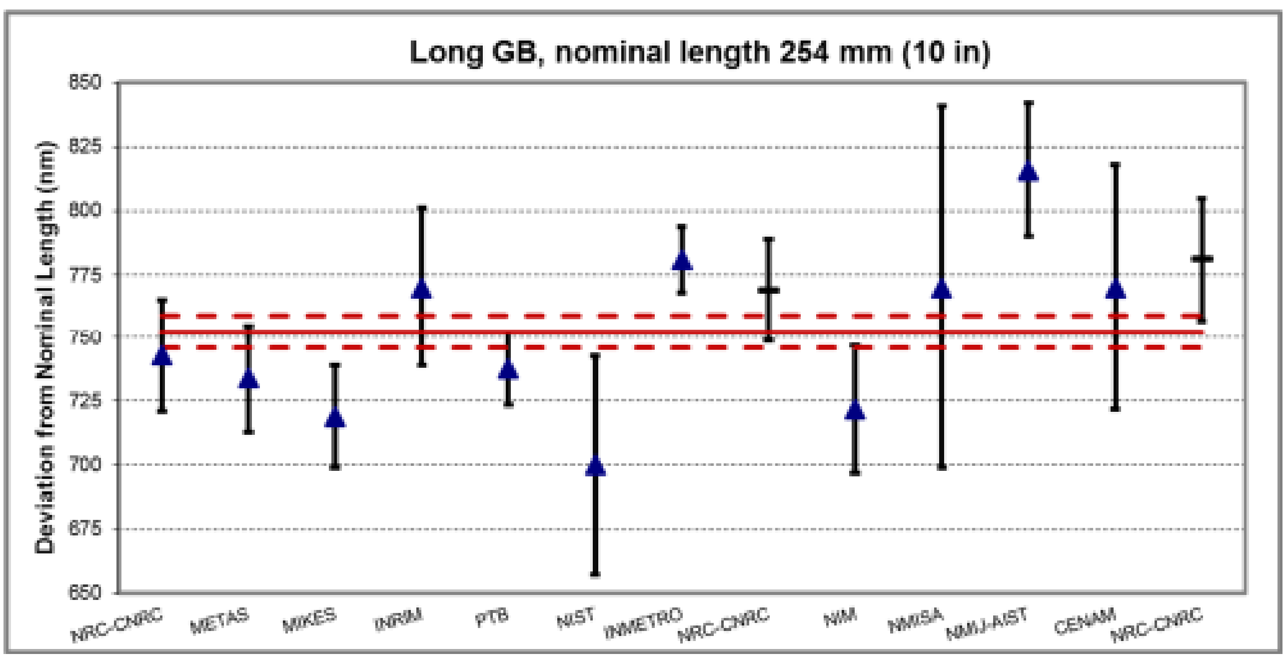

Figure 25 Measurement results of participants along with the weighted mean for the 254 mm long GB. All results are along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the upper and lower limit of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

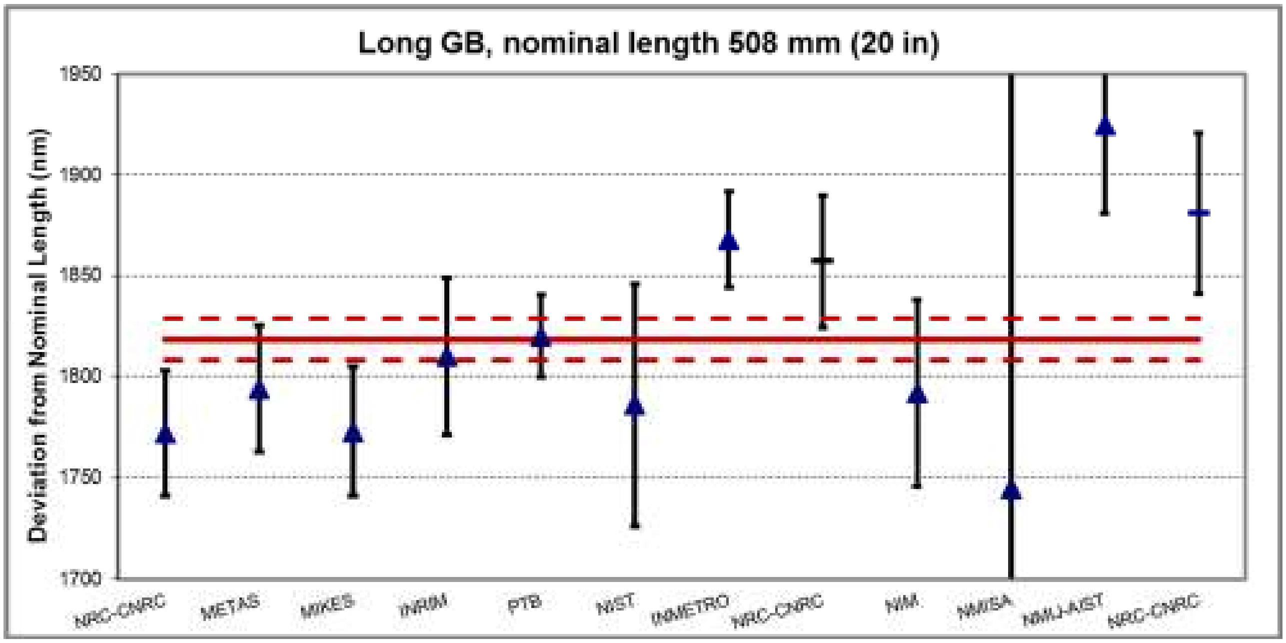

Figure 26 Measurement results of participants along with the weighted mean for the 508 mm long GB. All results are shown along with their standard uncertainty bars. The red solid line represents the weighted mean and the red dashed lines the upper and lower limit of its standard uncertainty. Values shown with a dash represent the second and third measurements by the pilot laboratory which were excluded from the calculation of the weighted mean.

8.2 Consideration of Drift

It was suggested that the 100 mm steel short gauge block as well as the three long gauge blocks could be subject to drift. Therefore, we performed a linear fit analysis on these gauge blocks to see whether drift should be considered.

The general equation to determine linear drift is:

$$\Delta l(t) = m(t - t_0) + b \tag{4}$$

where $\Delta l(t)$ is the variation in length over time t, m is the drift rate in nm/day, and b a length offset that depends on the choice of the time origin t_0.

To establish whether there is drift or not, the standard uncertainty of the slope, u_m, was calculated and compared to the value of the slope. If the absolute value of the slope $|m|$ is greater than its standard uncertainty, it is judged that there is drift. Otherwise, if $|m|$ is smaller or equal to its standard uncertainty, the no drift alternative ($m = 0$) is contained within the uncertainty interval and drift is not considered.

The idea was to determine the slope m only from the three pilot's control measurements at the beginning, middle and end of the circulation in order to have good reproducibility (same instrument, same phase change correction, etc.); while b, the offset, was determined by all the participants data. When t_0 is chosen as described below, b is actually equal to the weighted mean KCRV previously determined when drift is not taken into account. t_0 is defined as the central time or weighted mean of the measurement times. It is a time near the middle of the comparison time span.

w_i, the weighing factor of participant i, is defined as:

$$w_i = \left(\sum \frac{1}{\left({}^1\!/_{u_j} \right)^2} \right) \left({}^1\!/_{u_i} \right)^2 \tag{5}$$

Then the elapsed central time origin is given by: $t_0 = \sum w_i t_i$ (6)

Therefore, the elapsed time of laboratory i with respect to origin t_0 is: $t_i' = t_i - t_0$ (7)

8.2.1 100 mm Gauge Block

As the pilot laboratory, CENAM, spotted a problem on their instrument during the comparison exercise, as mentioned in **Appendix A**, their initial measurements were not reliable, and it was judged inadequate to determine the slope using this value. Additionally, the second and third measurements by the pilot laboratory were the same value. Therefore, for this specific case, we used all the participants' measurements to perform the linear fit excluding the first measurement by CENAM.

We obtain the following equation for the weighted mean as a function of time:

$$\bar{x}_w = -0.0078t' + 41.52 \quad \text{nm, } t' \text{ in days} \tag{8}$$

And an elapsed central time, t_0, corresponding to 2012-01-27.

The slope, m, -0.0078 nm/day is very small representing only -2.8 nm/year. Furthermore, the uncertainty of this slope, u_m, is 0.0087 nm/day. As $u_m \geq |m|$, the uncertainty of the slope is larger than the magnitude of the slope, therefore we cannot assert that there is drift. We conclude that drift should not be considered in this case. For completeness, we show the participants' results along with the weighted mean considering drift in figure **27**.

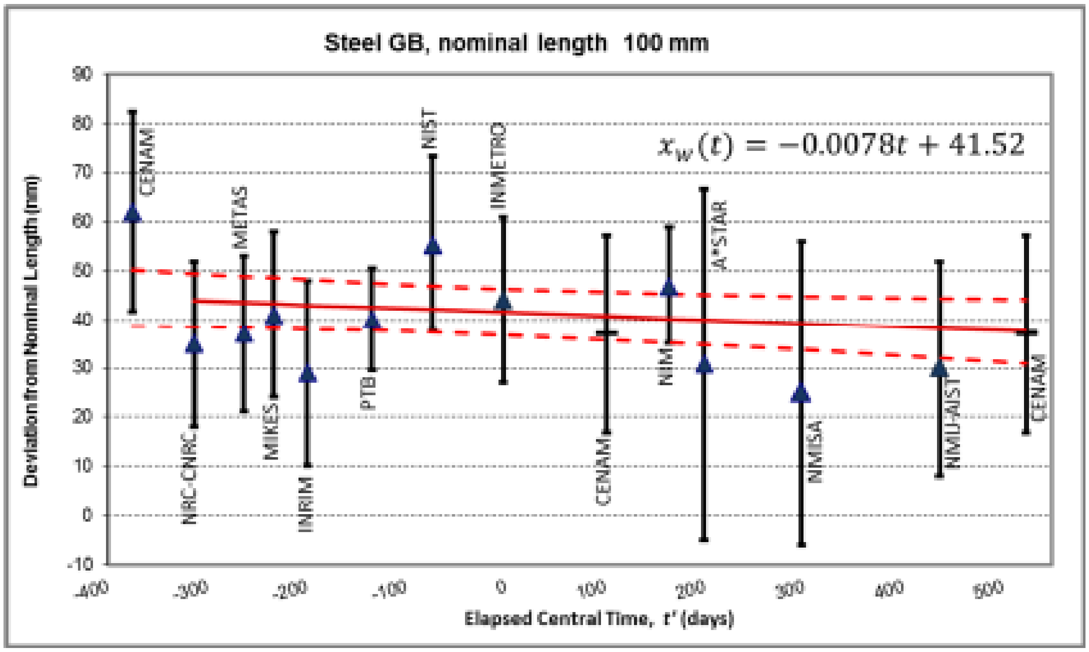

Figure 27 Drift analysis for the 100 mm gauge block. Drift slope calculated from all participants except the first measurement by CENAM. Ordinate to the origin calculated as the weighted mean of all participants depicted with triangles. The red solid line shows the weighted mean considering drift and the red dotted lines its standard uncertainty limits.

8.2.2 152.4 mm Gauge Block

A similar drift analysis was performed on the 152.4 mm gauge block. In this case, the pilot laboratory was NRC-CNRC and the three measurements are considered valid. Therefore, the slope was extracted from these three measurements. We obtained the following equation for the weighted mean as a function of time:

$$\bar{x}_w = -0.0152t' + 203.1 \quad \text{nm, } t' \text{ in days} \tag{9}$$

In this case we obtained a t_0 of 2012-03-23.

Figure **28** shows the participants' results along with the weighted mean as a function of time and its uncertainty.

We analyzed the drift in a similar way as for the 100 mm gauge block: the slope, m = -0.0152 nm/day (which represents -5.5 nm/year) is compared with its uncertainty, u_m of 0.02640 nm/day. As $u_m \geq |m|$, the uncertainty of the slope is larger than the magnitude of the slope, therefore we conclude we cannot say there is drift in this case neither and therefore, it was not considered.

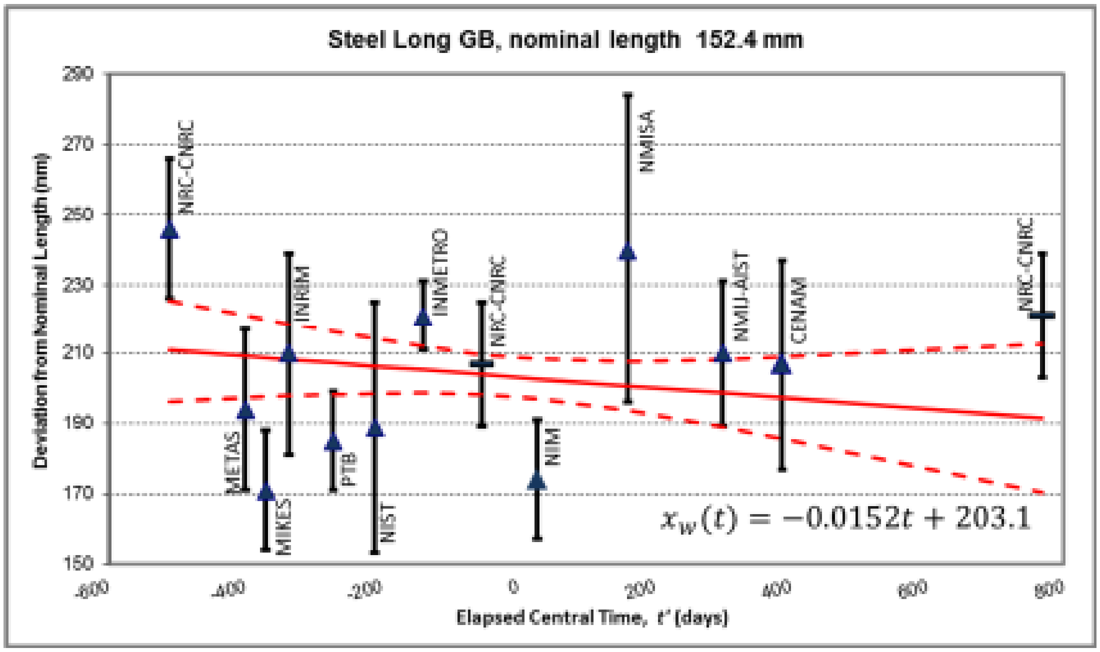

Figure 28 Drift analysis for the 152.4 mm gauge block. Drift slope calculated from the three pilot (NRC-CNRC) measurements. Ordinate to the origin is calculated as the weighted mean of all the participants data depicted with triangles. The red solid line shows the weighted mean considering drift and the red dotted lines its standard uncertainty limits.

8.2.3 254 mm Gauge Block

We performed the above drift analysis for the 254 mm gauge block as well. The three measurements from the pilot laboratory, NRC-CNRC, were used to determine the slope of the weighted mean. We obtained the following equation:

$$\bar{x}_w = 0.0279t' + 752.5 \qquad \text{nm, } t' \text{in days} \qquad (10)$$

For this gauge block t_0 corresponds to 2011-10-26.

In this case the gauge block grows at a rate of $m = 0.0279$ nm/day (which represents 10.2 nm/year). To judge the validity of this value we compare it again with its uncertainty, $u_m = 0.0109$ nm/day. In this case $u_m < |m|$ and drift may be confirmed. Therefore, drift should be considered in the subsequent analysis. Figure **29** shows the corresponding results.

8.2.4 508 mm Gauge Block

We repeated the drift analysis for the 508 mm gauge block. The three measurements from the pilot laboratory, NRC-CNRC, were used to determine the slope of the weighted mean. We obtained the following equation:

$$\bar{x}_w = 0.0785t' + 1818 \qquad \text{nm, } t' \text{ in days} \qquad (10)$$

For this gauge block t_0 corresponds to 2011-11-27.

Therefore, this gauge block grows at a rate of $m = 0.0785$ nm/day (which represents 28.7 nm/year). To judge the validity of this value, we compare it again with the uncertainty of the slope, $u_m = 0.0404$ nm/day. In this case $u_m < |m|$ again and drift may be confirmed. Therefore, drift should be considered in the subsequent analysis. Figure **30** shows the corresponding results.

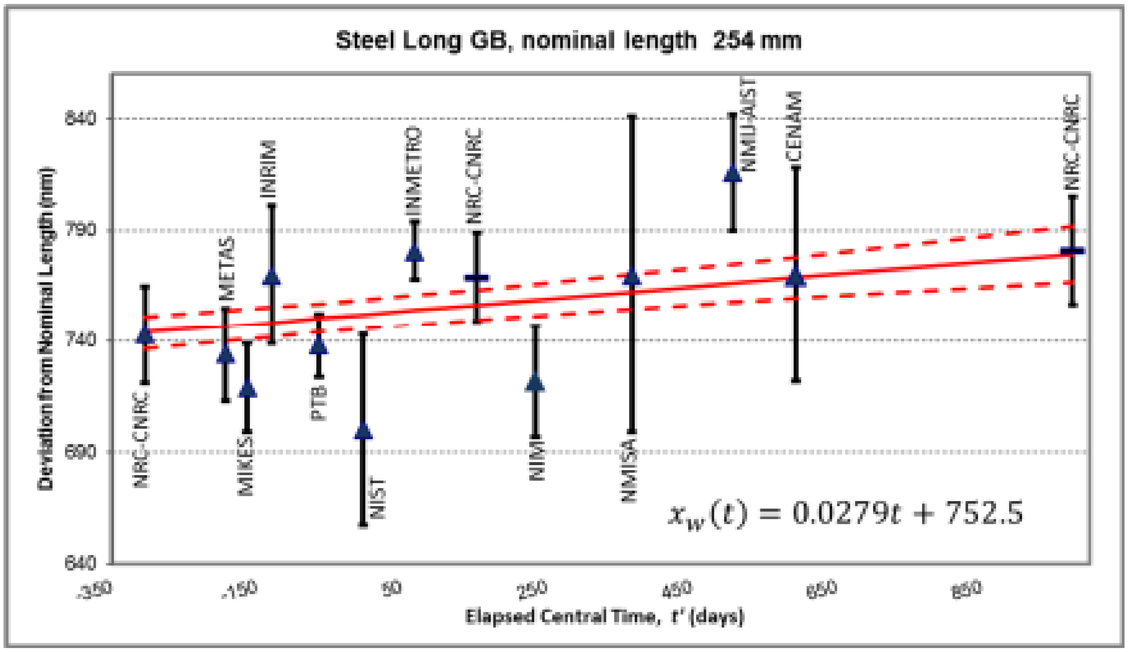

Figure 29 Drift analysis for the 254 mm gauge block. Drift slope calculated from the three pilot (NRC-CNRC) measurements. Ordinate to the origin calculated as the weighted mean of all participants depicted with triangles. The red solid line shows the weighted mean considering drift and the red dotted lines its standard uncertainty limits.

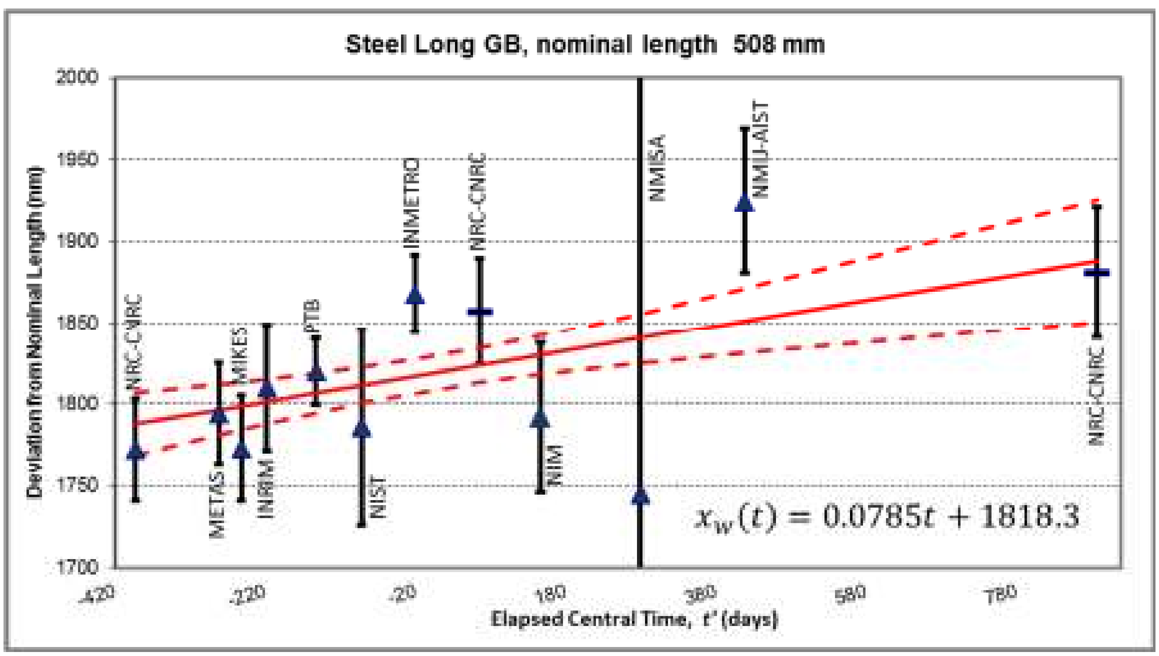

Figure 30 Drift analysis for the 508 mm gauge block. Drift slope calculated from the three pilot (NRC-CNRC) measurements. Ordinate to the origin calculated as the weighted mean of all participants depicted with triangles. The red solid line shows the weighted mean considering drift and the red dotted lines its standard uncertainty limits.

8.3 Normalized errors of the participant NMIs with respect to the Weighted Mean for each Gauge Block.

8.3.1 Definition of the normalized error

For the participants' results contributing to the determination of the weighted mean, the normalized errors, E_n, were calculated with the expanded uncertainties by means of the equation:

$$E_n = \frac{x_i - \bar{x}_W}{k_{95}\sqrt{u_{x_i}^2 - u_{\bar{x}_W}^2}}, \tag{11a}$$

and for the outliers not included in the determination of the weighted mean, the normalized errors are instead given by the following formula:

$$E_n = \frac{x_i - \bar{x}_W}{k_{95}\sqrt{u_{x_i}^2 + u_{\bar{x}_W}^2}}, \tag{11b}$$

Where, x_i, is the value obtained by laboratory i,
$\bar{x}_W$, is the weighted mean (or the time dependent weighted mean, $\bar{x}_W(t)$ when there is drift,
u_{x_i} and $u_{\bar{x}_W}$, are the standard uncertainties of, respectively, x_i and $\bar{x}_W$, and
k_{95} is the coverage factor for a confidence interval of 95 %. k_{95} depends on the effective degrees of freedom of x_i.

When $0 < E_n < 1$, it is judged that x_i is consistent with the weighted mean as reference value considering the expanded uncertainties.

Tables **16**, **17,** and **18** show the declared effective degrees of freedom (DOF), ν_{eff}, and the corresponding coverage factors for a 95% confidence interval[24], k_{95}, for each participant[25] for the short steel gauge blocks, the ceramics gauge blocks and the long steel gauge blocks respectively.

For this first analysis of the normalized errors, the calculation of the weighted mean considered all participants. In the case of the 254 mm gauge block and the 508 mm gauge block, drift was taken in account.

[24] Please note that these values are for a 95.00 % confidence level which differs slightly from the 95.45 % confidence level where k=2 corresponds to $\nu_{eff} = \infty$.

[25] If a participant did not declare the effective DOF, we considered $\nu_{eff} = 60$, with a coverage factor of $k_{95} = 2.00$.

Nom. Leng.	Steel Short Gauge Blocks																							
	CENAM		NRC-CNRC		METAS		MIKES		INRIM		PTB		NIST		INMETRO		NMIJ-AIST		NIM		A*STAR		NMISA	
mm	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$	ν_{eff}	$k_{95\%}$
0.5	42	2.02	10	2.23	51	2.01	32	2.04	445	1.97	314	1.97	60	2.00	10	2.23	∞	1.96	24	2.06	∞	1.96	-	-
3	42	2.02	10	2.23	54	2.00	33	2.03	447	1.97	218	1.97	60	2.00	10	2.23	∞	1.96	24	2.06	∞	1.96	35	2.03
5	43	2.02	10	2.23	12	2.18	33	2.03	452	1.97	219	1.97	60	2.00	10	2.23	∞	1.96	24	2.06	∞	1.96	35	2.03
7	44	2.02	10	2.23	55	2.00	33	2.03	458	1.97	219	1.97	60	2.00	10	2.23	∞	1.96	24	2.06	∞	1.96	35	2.03
25	63	2.00	10	2.23	65	2.00	41	2.02	388	1.97	234	1.97	60	2.00	15	2.13	∞	1.96	30	2.04	∞	1.96	57	2.00
80	184	1.97	10	2.23	151	1.98	89	1.99	128	1.98	345	1.97	60	2.00	30	2.04	∞	1.96	66	2.00	∞	1.96	568	1.96
100	189	1.97	10	2.23	372	1.97	92	1.99	50	2.01	376	1.97	60	2.00	43	2.02	∞	1.96	70	1.99	∞	1.96	1156	1.96

Table 16. Effective DOF, ν_{eff}, declared by the participating NMI and corresponding coverage factor for 95 % confidence interval, k_{95} for the Short Steel Gauge Blocks.

Nom. Leng.	Ceramics Gauge Blocks																							
	CENAM		NRC		METAS		MIKES		INRIM		PTB		NIST		INMETRO		NMIJ, AIST		NIM		A*STAR		NMISA	
mm	veff	k	veff	k	veff	k	veff	k	veff	k	veff	k	veff	k	veff	k	veff	k	veff	k	veff	k	veff	k
3	42	2.02	10	2.23	140	1.98	33	2.03	447	1.97	284	1.97	60	2.00	11	2.20	∞	1.96	24	2.06	∞	1.96	35	2.03
5	43	2.02	10	2.23	169	1.97	33	2.03	452	1.97	284	1.97	60	2.00	11	2.20	∞	1.96	24	2.06	∞	1.96	35	2.03
7	43	2.02	10	2.23	141	1.98	33	2.03	458	1.97	284	1.97	60	2.00	11	2.20	∞	1.96	24	2.06	∞	1.96	35	2.03
10	45	2.01	10	2.23	139	1.98	34	2.03	468	1.97	285	1.97	60	2.00	11	2.20	∞	1.96	25	2.06	∞	1.96	37	2.03
80	182	1.97	10	2.23	259	1.97	89	1.99	62	2.00	339	1.97	60	2.00	29	2.05	∞	1.96	66	2.00	∞	1.96	568	1.96
90	201	1.97	10	2.23	331	1.97	91	1.99	55	2.00	367	1.97	60	2.00	34	2.03	∞	1.96	68	2.00	∞	1.96	820	1.96

Table 17. Effective DOF, ν_{eff}, declared by the participating NMI and corresponding coverage factor for 95 % confidence interval, k_{95} for the Ceramics Gauge Blocks.

Nom. Leng.	Long Gauge Blocks																					
	CENAM		NRC-CNRC		METAS		MIKES		INRIM		PTB		NIST		INMETRO		NMIJ-AIST		NIM		NMISA	
mm	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k	ν_{eff}	k
152.4	231	1.97	10	2.23	38	2.02	63	2.00	124	1.98	229	1.97	60	2.00	28	2.05	∞	1.96	65	2.00	5041	1.96
254	151	1.98	10	2.23	113	1.98	123	1.98	150	1.98	330	1.97	60	2.00	57	2.00	∞	1.96	56	2.00	35000	1.96
508	-	-	10	2.23	313	1.97	680	1.96	170	1.97	421	1.97	60	2.00	68	2.00	∞	1.96	45	2.01	33	2.03

Table 18. Effective DOF, ν_{eff}, declared by the participating NMI and corresponding coverage factor for 95 % confidence interval, k_{95} for the Long Steel Gauge Blocks.

Table **19** shows the normalized errors calculated with equation **11a** for the steel short gauge blocks from the participating NMIs; table **20** for the six ceramic gauge blocks; and table **21** for the three steel long gauge blocks. In all cases we considered in this first calculation the results of all participants to determine the weighted mean. Drift was considered for the cases of the 254 mm and 508 mm gauge blocks according to the conclusions derived from the analysis in section 8.2.

Steel Short Gauge Blocks							
Nominal Length (mm)	**0.5**	**3**	**5**	**7**	**25**	**80**	**100**
NMI	**Normalized Errors, E_n**						
A*STAR	0.17	0.03	-0.10	1.13	0.07	-0.21	-0.15
CENAM	0.22	0.37	1.41	1.08	1.28	0.72	0.53
INMETRO	-0.43	0.33	-0.03	-0.48	0.01	0.62	0.08
INRIM	0.05	0.04	-0.46	0.21	-0.25	-0.54	-0.34
METAS	-0.58	-0.38	0.55	-0.60	-1.25	-0.14	-0.15
MIKES	0.41	-0.12	-0.19	-0.18	0.25	-0.15	-0.02
NIM	-0.40	-0.41	0.25	0.34	0.73	0.37	0.25
NIST	1.26	1.20	0.77	0.49	0.31	0.03	0.40
NMIJ-AIST	-	-0.34	-0.39	-0.26	-0.55	-0.41	-0.27
NMISA	-	0.72	-0.34	-0.37	-0.12	-0.11	-0.27
NRC-CNRC	-0.12	-0.44	-0.30	-0.35	0.16	0.14	-0.18
PTB	-0.57	-0.85	-1.26	-1.05	-0.81	-0.44	-0.08

Table 19. Normalized errors for the seven steel short GB. The weighted mean calculation considered the contribution of all NMIs.

Ceramic Short Gauge Blocks						
Nominal Length (mm)	**3**	**5**	**7**	**10**	**80**	**90**
NMI	**Normalized Errors, E_n**					
A*STAR	-0.19	-0.26	-0.13	0.20	-0.95	-0.93
CENAM	1.10	0.36	0.84	0.32	-0.29	0.08
INMETRO	0.46	0.68	-0.63	0.70	0.61	0.99
INRIM	-0.26	-0.03	-0.12	-0.25	-0.11	-0.50
METAS	0.03	0.78	0.60	0.31	0.59	0.39
MIKES	0.24	-0.13	0.47	0.30	0.03	0.19
NIM	-1.03	-1.04	-0.46	-0.56	-1.24	-1.33
NIST	0.21	0.14	0.00	-0.23	-0.27	0.08
NMIJ-AIST	0.63	0.67	0.57	0.44	0.56	0.31
NMISA	-0.99	-1.00	-1.25	-0.75	0.40	0.36
NRC-CNRC	0.38	0.44	0.55	0.36	0.59	0.52
PTB	-0.38	-0.13	-0.37	-0.57	0.20	-0.08

Table 20. Normalized Errors for the six Ceramic short GB. Weighted mean calculation considering the contribution of all NMIs.

Steel Long Gauge Blocks			
Nom. Length (mm)	**152.4**	**254**	**508**
NMI	**Normalized Errors, E_n**		
CENAM	0.07	-0.01	-
INMETRO	1.05	1.18	0.92
INRIM	0.12	0.35	-0.02
METAS	-0.20	-0.33	-0.22
MIKES	-1.00	-0.77	-0.60
NIM	-0.91	-0.76	-0.52
NIST	-0.20	-0.62	-0.30
NMIJ-AIST	0.17	0.98	0.67
NMISA	0.43	0.05	-0.16
NRC-CNRC I	1.00	-0.02	-0.43
PTB	-0.72	-0.50	0.09

Table 21. Normalized Errors for the three steel long gauge blocks. Drift analysis was considered for the 254 mm and the 508 mm GB. The weighted mean was calculated considering the contribution of all NMIs.

8.4 Exclusion of outliers

Observing the E_n values from Tables **19**, **20** and **21** we see that a few were greater than 1. These values were considered as outliers. We proceeded to withdraw the largest value for each gauge block, and we recalculated a new weighted mean without it. If necessary, the process was repeated until the E_n values of the remaining gauge blocks were all under one.

8.4.1 First Iteration

In a first iteration we excluded the NMIs that had the largest E_n values from the corresponding weighted mean calculation. For the short steel gauge blocks we removed:

- NIST from the 0.5 mm,
- NIST from the 3 mm,
- CENAM from the 5 mm,
- A*STAR from the 7 mm and
- CENAM from the 25 mm.

For the ceramic gauge blocks we withdrew:

- CENAM from the 3 mm,
- NIM from the 5 mm,
- NMISA from the 7 mm and
- NIM from the 80 mm and
- NIM from the 90 mm

For the long steel gauge blocks we withdrew:

- INMETRO from the 152.4 and
- INMETRO from the 254.

In addition, NMISA was excluded from the calculation of the reference value for the 254 mm and the 508 mm gauge blocks because the quoted uncertainties were greater than the required uncertainty established in the protocol and quoted in Section 2.1.

We recalculated the new weighted means considering the elimination of the largest E_n values for those gauge blocks where there were values greater than one. This was required for twelve cases that are shown in Table **22**.

Nominal Length (mm)	Weighted Mean (nm)		Nominal Length (mm)	Weighted Mean (nm)	
	$\bar{x}_W$	$u(\bar{x}_W)$		$\bar{x}_W$	$u(\bar{x}_W)$
0.5 mm steel	-23.8	3.3	5 mm ceramic	39.8	3.5
3 mm steel	-31.0	3.1	7 mm ceramic	46.9	3.2
5 mm steel	11.8	3.2	80 mm ceramic	56.7	4.9
7 mm steel	6.2	3.0	90 mm ceramic	53.3	5.2
25 mm steel	9.2	3.2	152.4 mm long	195.1	6.7
3 mm ceramic	-19.1	3.3	254 mm long	742.2[26]	7.7

Table 22. Weighted Mean after a first iteration of elimination of outliers from the calculation.

Tables **23a** and **23b** show the new E_n values after this first iteration where we eliminated the outliers with previously largest E_n values from the calculation of the weighted mean for the twelve gauge blocks considered.

Gauge Block	0.5 mm Steel	3 mm Steel	5 mm Steel	7 mm Steel	25 mm Steel	3 mm Ceramic
NMI	**Normalized Errors, E_n**					
A*STAR	0.28	0.12	-0.00	1.13	0.14	-0.10
CENAM	0.39	0.50	1.41	1.16	1.28	1.10
INMETRO	-0.32	0.42	0.09	-0.43	0.10	0.55
INRIM	0.23	0.18	-0.29	0.29	-0.12	-0.12
METAS	-0.43	-0.26	0.63	-0.53	-1.12	0.12
MIKES	0.56	-0.00	-0.04	-0.11	0.37	0.37
NIM	-0.17	-0.23	0.48	0.45	0.93	-0.86
NIST	1.26	1.20	0.94	0.57	0.43	0.35
NMIJ-AIST	-	-0.24	-0.28	-0.22	-0.45	0.73
NMISA	-	0.82	-0.23	-0.32	-0.04	-0.90
NRC-CNRC	-0.05	-0.37	-0.21	-0.31	0.24	0.46
PTB	-0.41	-0.71	-1.08	-0.97	-0.66	-0.30

Table 23a. E_n values obtained after a first elimination of outliers for the considered gauge blocks. Highlighted in yellow are shown the values that were eliminated from the calculation.

[26] Drift is considered.

Gauge Block	5 mm Ceramic	7 mm Ceramic	80 mm Ceramic	90 mm Ceramic	152.4 mm Steel	254 mm Steel
NMI	**Normalized Errors, E_n**					
A*STAR	-0.38	-0.20	-1.05	-1.04	-	-
CENAM	0.18	0.74	-0.47	-0.10	0.07	0.11
INMETRO	0.56	-0.70	0.43	0.80	0.72	1.18
INRIM	-0.17	-0.23	-0.29	-0.67	0.12	0.51
METAS	0.65	0.53	0.41	0.19	-0.21	-0.10
MIKES	-0.30	0.37	-0.17	-0.01	-1.03	-0.54
NIM	-1.04	-0.61	-1.24	-1.33	-0.93	-0.58
NIST	-0.05	-0.10	-0.45	-0.10	-0.20	-0.51
NMIJ	0.54	0.49	0.37	0.10	0.18	1.17
NMISA	-1.13	-1.25	0.30	0.25	0.43	0.12
NRC-CNRC	0.34	0.49	0.42	0.34	1.02	0.18
PTB	-0.26	-0.44	-0.01	-0.30	-0.75	-0.14

Table 23b. E_n values obtained after a first elimination of outliers for the considered gauge blocks. Highlighted in yellow are shown the values that were eliminated from the calculation.

8.4.2 Second Iteration

We may observe from tables **23a** and **23b** that there are still other values greater than one on eight gauge blocks. We continued eliminating the NMIs that had E_n values greater than one from the corresponding weighted mean calculation. For the short steel gauge blocks we still removed:

- PTB from the 5 mm,
- CENAM from the 7 mm and
- METAS from the 25 mm.

For the ceramic gauge blocks we withdrew additionally:

- NMISA from the 5 mm,
- A*STAR from the 80 mm and
- A*STAR from the 90 mm

For the long steel gauge blocks we still withdrew:

- NRC-CNRC from the 152.4 and
- NMIJ-AIST from the 254.

We proceed to eliminate them from the calculation of the reference value in a second iteration. We obtain the following new values:

Nominal Length (mm)	Weigthed Mean (nm)		Nominal Length (mm)	Weigthed Mean (nm)	
	$\bar{x}_W$	$u(\bar{x}_W)$		$\bar{x}_W$	$u(\bar{x}_W)$
5 mm steel	14.5	3.4	**80 mm ceramic**	58.5	3.4
7 mm steel	3.8	3.2	**90 mm ceramic**	55.1	5.3
25 mm steel	11.7	3.4	**152.4 mm long**	188.7	7.1
5 mm ceramic	42.0	3.6	**254 mm long**	735.0[27]	8.1

Table 24. Weighted Mean after elimination of outliers on a second iteration.

Table **25** shows the new E_n values after the second iteration where we eliminated the outliers with previously largest E_n values from the calculation of the weighted mean for the eight remaining gauge blocks.

Gauge Block	5 mm Steel	7 mm Steel	25 mm Steel	5 mm ceramic	80 mm ceramic	90 mm ceramic	152.4 mm long steel	254 mm long steel
NMI	**Normalized Errors, E_n**							
A*STAR	-0.10	1.21	0.07	-0.46	-1.05	-1.04	-	-
CENAM	1.26	1.16	1.16	0.06	-0.53	-0.15	0.32	0.17
INMETRO	-0.02	-0.34	0.01	0.48	0.36	0.75	1.29	1.35
INRIM	-0.46	0.44	-0.25	-0.31	-0.35	-0.73	0.38	0.61
METAS	0.56	-0.41	-1.12	0.57	0.34	0.13	0.12	0.05
MIKES	-0.18	-0.01	0.25	-0.42	-0.23	-0.07	-0.57	-0.39
NIM	0.28	0.64	0.75	-1.17	-1.32	-1.40	-0.48	-0.46
NIST	0.79	0.71	0.31	-0.17	-0.51	-0.16	0.00	-0.45
NMIJ-AIST	-0.39	-0.16	-0.56	0.45	0.31	0.03	0.55	1.17
NMISA	-0.34	-0.23	-0.13	-1.13	0.27	0.22	0.60	0.16[28]
NRC-CNRC	-0.29	-0.24	0.16	0.28	0.36	0.29	1.21	0.31
PTB	-1.08	-0.82	-0.83	-0.34	-0.05	-0.37	-0.16	0.09

Table 25. E_n values obtained after the second elimination of outliers for the eight gauge blocks considered. Highlighted in yellow are shown the values that were eliminated from the calculation in the first iteration and on magenta those eliminated in this second iteration.

The elimination process was halted after the second iteration as all the remaining E_n values were smaller than one. The corresponding Key Comparison Reference Values (KCRV) for the whole set of gauge blocks is summarized in Table **26**.

[27] Drift is considered.
[28] This value was also excluded from the calculation of the KCRV as its uncertainty exceeded the limits established in section 2.1.

Nominal Length (mm)	KCRV - Steel Short Gauge Blocks (nm)		Nominal Length (mm)	KCRV - Ceramic Gauge Blocks (nm)		Nominal Length (mm)	KCRV - Steel Long Gauge Blocks (nm)	
	$\bar{x}_W$	$u(\bar{x}_W)$		$\bar{x}_W$	$u(\bar{x}_W)$		$\bar{x}_W$	$u(\bar{x}_W)$
0.5	-23.8	3.3	3	-19.1	3.3	152.4	188.7	7.1
3	-31.0	3.1	5	42.0	3.6	254	735.0[29]	8.1
5	14.5	3.4	7	46.9	3.2	508	1818.4[30]	10.5
7	3.8	3.2	10	25.2	3.1			
25	11.7	3.4	80	58.5	5.0			
80	27.4	4.3	90	55.1	5.3			
100	41.5	4.9						

Table 26. Final Key Comparison Reference Values (KCRV) for all the gauge blocks.

8.5 Consistency of Declared Uncertainties with respect to the comparison exercise

u_{int}, the internal standard uncertainty of the KCRV is defined in [6], and is the combined standard uncertainty of the weighted mean defined in (2):

$$u_{int} = 1 \Big/ \sqrt{\sum \left(\frac{1}{u_i}\right)^2} \qquad (12)$$

Where, u_i, is the standard uncertainty of laboratory i.

And u_{ext}, the external uncertainty of the KCRV, defined as the standard deviation of the set of participants results [6]:

$$u_{ext} = \sqrt{\frac{\sum \frac{1}{u_i^2}(x_i - \bar{x}_W)^2}{(n-1)\sum \frac{1}{u_i^2}}} \qquad (13)$$

Where, x_i, is the result of laboratory i,
$\bar{x}_W$, is the KCRV and
n, the number of participants.

Then the Birge ratio is defined as: $$R_B = {}^{u_{ext}} \big/ {}_{u_{int}} \qquad (14)$$

[29] Drift is considered.
[30] Drift is considered.

The data is considered consistent with a coverage factor of 95% provided that [6],

$$R_B < \sqrt{1 + \sqrt{8/(n-1)}} \tag{15}$$

The following **Tables**, **27**, **28** and **29**, show the Birge ratios obtained for each of the comparison artifacts for the final results considering the removal of outliers.

Birge Ratio, Steel Gauge Blocks					
Nominal Length (mm)	u_{int}	u_{ext}	R_{Birge}	n	*Smaller than*
0.5	3.33	2.35	0.71	9	1.41
3	3.07	2.66	0.87	11	1.38
5	3.43	2.82	0.82	10	1.39
7	3.20	2.97	0.93	10	1.39
25	3.43	2.83	0.82	10	1.39
80	4.29	3.31	0.77	12	1.36
100	4.85	2.61	0.54	12	1.36

Table 27. Internal and external uncertainties of the KCRV, the Birge ratio, the number of contributions to the KCRV and the corresponding criterion (Eq. 15) to determine if the set of values are statistically consistent for the steel short gauge blocks.

Birge Ratio, Ceramic Gauge Blocks					
Nominal Length (mm)	u_{int}	u_{ext}	R_{Birge}	n	*Smaller than*
3	3.26	3.46	1.06	11	1.38
5	3.58	2.80	0.78	10	1.39
7	3.19	3.20	1.00	11	1.38
10	3.11	2.90	0.93	12	1.36
80	5.03	3.60	0.72	10	1.39
90	5.27	5.21	0.99	10	1.39

Table 28. Internal and external uncertainties of the KCRV, the Birge ratio, the number of contributions to the KCRV and the corresponding criterion (Eq.15) to determine if the set of values are statistically consistent for the ceramic gauge blocks.

Birge Ratio, Long Gauge Blocks					
Nominal Length (mm)	u_{int}	u_{ext}	R_{Birge}	*n*	***Smaller than***
152.4	7.11	5.57	0.78	9	1.41
254	8.11	6.00	0.74	8	1.44
508	10.48	10.94	1.04	9	1.41

Table 29 Internal and external uncertainties of the KCRV, the Birge ratio, the number of contributions to the KCRV and the corresponding criterion (Eq. 15) to determine if the set of values are statistically consistent for the long steel Gauge Blocks.

For the whole set of sixteen artifacts, the set of declared uncertainties was consistent with the results obtained in the comparison exercise, as may be seen from the above tables. In all cases the Birge ratio is smaller than the maximum value required for statistical significance for a 95% coverage factor.

Furthermore, the results are also statistically consistent when no outliers are removed from the measurement sets. This is shown in Tables **30**, **31** and **32** where all the measurements satisfy the criteria of equation 15.

Birge Ratio, Steel Gauge Blocks					
Nominal Length (mm)	u_{int}	u_{ext}	R_{Birge}	*n*	***Smaller than***
0.5	3.13	3.35	1.07	10	1.39
3	2.91	3.20	1.10	12	1.36
5	3.01	3.93	1.31	12	1.36
7	2.96	3.81	1.29	12	1.36
25	3.10	3.93	1.27	12	1.36
80	4.29	3.31	0.77	12	1.36
100	4.85	2.61	0.54	12	1.36

Table 30. Internal and external uncertainties of the KCRV, the Birge ratio, the number of contributions to the KCRV without the removal of outliers and the corresponding criterion (Eq. 15) to determine if the values are statistically consistent for the steel short gauge blocks.

Birge Ratio, Ceramic Gauge Blocks					
Nominal Length (mm)	u_{int}	u_{ext}	R_{Birge}	*n*	***Smaller than***
3	3.08	3.73	1.21	12	1.36
5	3.10	3.66	1.18	12	1.36
7	3.10	3.79	1.22	12	1.36
10	3.11	2.90	0.93	12	1.36
80	4.43	5.18	1.17	12	1.36
90	4.70	5.73	1.22	12	1.36

Table 31. Internal and external uncertainties of the KCRV, the Birge ratio, the number of contributions to the KCRV without the removal of outliers and the corresponding criterion (Eq.15) to determine if the values are statistically consistent for the ceramic gauge blocks.

Birge Ratio, Long Gauge Blocks					
Nominal Length (mm)	u_{int}	u_{ext}	R_{Birge}	*n*	*Smaller than*
152.4	5.57	7.29	1.31	11	1.38
254	6.62	8.17	1.23	11	1.38
508	10.48	10.30	0.98	10	1.39

Table 32 Internal and external uncertainties of the KCRV, the Birge ratio, the number of contributions to the KCRV without the removal of outliers and the corresponding criterion (Eq. 15) to determine if of values are statistically consistent for the long steel Gauge Blocks.

9 Final Results Summary

Tables, **33**, **34** and **35** summarize the final E_n values obtained by the participants for all artifacts of the short steel Gauge Blocks, ceramic Gauge Blocks, and long steel Gauge Blocks respectively.

Steel Short Gauge Blocks							
Nominal Length (mm)	0.5	3	5	7	25	80	100
NMI	Normalized Errors, E_n						
A*STAR	0.28	0.12	-0.10	1.21	0.07	-0.21	-0.15
CENAM	0.39	0.50	1.26	1.16	1.16	0.72	0.53
INMETRO	-0.32	0.42	-0.02	-0.34	0.01	0.62	0.08
INRIM	0.23	0.18	-0.46	0.44	-0.25	-0.54	-0.34
METAS	-0.43	-0.26	0.56	-0.41	-1.12	-0.14	-0.15
MIKES	0.56	-0.00	-0.18	-0.01	0.25	-0.15	-0.02
NIM	-0.17	-0.23	0.28	0.64	0.75	0.37	0.25
NIST	1.26	1.20	0.79	0.71	0.31	0.03	0.40
NMIJ-AIST	-	-0.24	-0.39	-0.16	-0.56	-0.41	-0.27
NMISA	-	0.82	-0.34	-0.23	-0.13	-0.11	-0.27
NRC-CNRC	-0.05	-0.37	-0.29	-0.24	0.16	0.14	-0.18
PTB	-0.41	-0.71	-1.08	-0.82	-0.83	-0.44	-0.08

Table 33. Final normalized errors obtained for the steel short gauge blocks.

Ceramic Short Gauge Blocks						
Nominal Length (mm)	**3**	**5**	**7**	**10**	**80**	**90**
NMI	**Normalized Errors, E_n**					
A*STAR	-0.10	-0.46	-0.20	0.20	-1.05	-1.04
CENAM	1.10	0.06	0.74	0.32	-0.53	-0.15
INMETRO	0.55	0.48	-0.70	0.70	0.36	0.75
INRIM	-0.12	-0.31	-0.23	-0.25	-0.35	-0.73
METAS	0.12	0.57	0.53	0.31	0.34	0.13
MIKES	0.37	-0.42	0.37	0.30	-0.23	-0.07
NIM	-0.86	-1.17	-0.61	-0.56	-1.32	-1.40
NIST	0.35	-0.17	-0.10	-0.23	-0.51	-0.16
NMIJ-AIST	0.73	0.45	0.49	0.44	0.31	0.03
NMISA	-0.90	-1.13	-1.25	-0.75	0.27	0.22
NRC-CNRC	0.46	0.28	0.49	0.36	0.36	0.29
PTB	-0.30	-0.34	-0.44	-0.57	-0.05	-0.37

Table 34. Final normalized errors obtained for the ceramic gauge blocks.

Steel Long Gauge Blocks			
Nom. Length (mm)	**152.4**	**254**	**508**
NMI	**Normalized Errors, E_n**		
CENAM	0.32	0.17	-
INMETRO	1.29	1.35	0.92
INRIM	0.38	0.61	-0.02
METAS	0.12	0.05	-0.22
MIKES	-0.57	-0.39	-0.60
NIM	-0.48	-0.46	-0.52
NIST	0.00	-0.45	-0.30
NMIJ-AIST	0.55	1.17	0.67
NMISA	0.60	0.16[31]	-0.16
NRC-CNRC I	1.21	0.31	-0.43
PTB	-0.16	0.09	0.09

Table 35. Final normalized errors obtained for the steel long gauge blocks. Drift analysis was considered for the 254 mm and the 508 mm GB's.

10 Conclusions

- The comparison exercise was successful, the artifacts behaved adequately throughout the comparison and no appreciable change was detected apart from the 254 mm and the 508 mm GB where drifts were detected. These two gauge blocks grew at a rate of 10.2 nm/year and 28.7 nm/year, respectively.
- The number of participants was ideal for a CCL key comparison. Twelve is a sufficiently large number to give reliable statistical numbers but it is not too large to make the circulation of the artefacts too long or for the artefacts to get completely damaged.

[31] This value was also excluded from the calculation of the KCRV as its uncertainty exceeded the limits established in section 2.1.

- The GBs came back to the pilot laboratories with damage more severe than expected but they could all be wrung and remeasured at the end of the comparison.
- Although there were delays in the circulation, the comparison was successfully completed.
- Drift was not initially considered by the pilots. However, it was pointed out by several participants that it could be present especially in the long GBs[32]. Drift was found for the 254 mm and the 508 mm blocks. It was estimated and considered for the comparison calculations. Future comparisons of long GBs should always consider the possibility of drift being present.
- For short GBs, all laboratories applied absolute interferometry fringe fraction evaluation. The instruments used in the key comparison were five commercial interferometers NPL-TESA, two NRLM-TSUGAMI, one ZEISS-Jena, one MITUTOYO, and one homemade Twyman-Green interferometer. Additionally, two Hilger interferometers and one ZEISS-Jena Koesters interferometer were also used. Most of the commercial models were modified to apply image processing and some of them phase shifting. Most of them used two or more optical wavelengths. As the techniques applied were similar, the uncertainties were of the same order of magnitude.
- For the long GBs, the laboratories used different techniques and arrangements. Some laboratories measured the gauge blocks in the vertical position while most of them measured them in the horizontal position. One laboratory used a CMM coupled to a heterodyne interferometer. Therefore, the uncertainties have a broader spread. This was the first time that such a technique was used to measure GBs in a key comparison. The results were remarkably good.
- At the 2016 working Group in dimensional Metrology of CCL meeting at the Dutch Metrology Institute (VSL) in Delft, the weighted mean was chosen for the calculation of the KCRV. This is recommended for all other future comparisons.
- From the observation of the results in the graphs and the calculated E_n values, most of the results were in good agreement with the KCRV considering their claimed uncertainties. These declared uncertainties were also reasonable when judging the performance from the Birge ratios shown in Tables 27, 28, and 29 for the final results. In addition, the same analysis was performed without the removal of outliers as shown in Tables 30, 31, and 32. In all cases the Birge ratios indicate statistical consistency.
- In the opinion of the pilots, the results obtained are technically valid to support the participants' Calibration and Measurement Capabilities (CMCs) on GB Calibration by Optical Interferometry.
- In the case of CENAM, which had four E_n values greater than 1.0, the problem was identified and solved during the comparison as described in Problem 1 of **Appendix A**. We believe the corroboration of the effectiveness of the solution are the closure results obtained in this comparison by CENAM.

11. References.

1. ISO 3650:1998(E), *Geometrical Product Specification (GPS) – Length Standards – Gauge Blocks*, International Organization for Standardization, Geneva, Switzerland.
2. *Guide to the Expression of Uncertainty in Measurement,* International Organization for Standardization, 1995, 110p., Geneva, Switzerland.

[32] We specially acknowledge this observation by NMIJ-AIST from Japan.

3. J. E. Decker, J. R. Pekelsky, *Uncertainty evaluation for the measurement of gauge blocks by optical interferometry*, *Metrologia*, **34**, 479-493 (1997).
4. J. E. Decker, N. Brown *et al.*, *Recent recommendations of Consultative Committee for Length (CCL) regarding strategies for evaluating key comparison data. Metrologia* **43** (2006) L51-L55.
5. R. Thalmann, CCL-K1 *Calibration of Gauge Blocks by Interferometry Comparison Report*, Wabern, November 2000.
6. R. Thalmann, CCL key comparison: calibration of gauge blocks by interferometry, Metrologia **39**, 165 (2002)

Appendix A Problems experienced by the pilot laboratory for short GB measurements during the comparison

Problem 1.
Juan Carlos Zárraga and Carlos Colín measured the GBs circulated in this comparison in February and March 2011. The platen used had been relapped recently at NPL in November 2010. The phase change corrections (PCC) calculated by the two operators were different. Carlos Colín tends to give a PCC always about -15 nm from that of Juan Carlos. However, this difference also depends on the platen and we had no previous information on this platen by either of the operators after it was relapped. The results sent to Andrew Lewis, who acted as escrow for CENAM, were those of Juan Carlos. However, by measuring the PCC afterwards with this platen for well-known GBs, we realized the right value should have been -55 nm instead of –35 nm for this platen.

Problem 2.
We experienced a second problem that was detected in July 2013. While measuring a GB set from a customer, a relatively large dispersion of values was detected between repeated measurements. The variations were around 40 nm for a 25 mm GB and, apparently, it was length dependent. The first things checked were the temperature measurement, the temperature probes and temperature stability. The sensors were immediately verified and we reviewed our temperature records. Everything was under control and a temperature problem was discarded. Then we suspected the two lasers, green and red, that could have had a frequency change. As we knew the green laser was prone to changing mode, the measurement was performed with the red laser only. Nevertheless, the same variations in length persisted. Then we changed the red laser for a recently calibrated one. The problem was not solved.

Afterwards, all the peripheral additional measurements were checked. The standard resistor for the temperature bridge was changed for a spare one, we borrowed a barometer from our colleagues in the Pressure Section. The readings sensibly were the same in both cases.

At the end, we checked the humidity dew-point meter against another humidity meter and differences were registered. The due-point meter was then withdrawn and opened. The diaphragm that takes the air sample was ripped. We concluded the air sample was not taken properly. The original MITCHELL humidity meter had been recently replaced with this EDGE meter that was damaged. We then reinstalled the original MITCHELL instrument. The variation in length was reduced but was still present.

The Thermometry Section had recently calibrated this instrument. We reviewed the calibration certificate and we realized that the electric tension at which it was calibrated, was very different from the values of the previous calibrations and therefore, large differences were reported for the humidity parameters.

We immediately introduced the original manufacturers calibration parameters to the interferometer software. Finally, the variations disappeared. Then, we requested our colleagues from the Thermometry Section to recalibrate it with the prescribed tension values. These values were corrected in the interferometer software and the problem was finally solved in March 2014.