NRC Publications Archive
Archives des publications du CNRC

A comparison of future realizations of the kilogram

Stock, Michael; Barat, Pauline; Pinot, Patrick; Beaudoux, Florian; Espel, Patrick; Piquemal, Francois; Thomas, Matthieu; Ziane, Djamel; Abbott, Patrick; Haddad, Darine; Zubarych, Zeina; Pratt, Jon R.; Schlamminger, Stephan; Fujii, Kenichi; Fujita, Kazuaki; Kuramoto, Naoki; Mizushima, Shigeki; Zhang, Lulu; Davidson, Stuart; Green, Richard; Liard, Jacques; Sanchez, Carlos; Wood, Barry; Bettin, Horst; Borys, Michael; Busch, Ingo; Hämpke, Michael; Krumrey, Michael; Nicolaus, Arnold R.

This publication could be one of several versions: author’s original, accepted manuscript or the publisher’s version. / La version de cette publication peut être l’une des suivantes : la version prépublication de l’auteur, la version acceptée du manuscrit ou la version de l’éditeur.
For the publisher’s version, please access the DOI link below./ Pour consulter la version de l’éditeur, utilisez le lien DOI ci-dessous.

Publisher’s version / Version de l'éditeur:
https://doi.org/10.1088/1681-7575/aa9a7e
Metrologia, 2017-11-14

NRC Publications Record / Notice d'Archives des publications de CNRC:
https://nrc-publications.canada.ca/eng/view/object/?id=1d721f83-1467-4fb7-93ad-be9dc3dcd41d
https://publications-cnrc.canada.ca/fra/voir/objet/?id=1d721f83-1467-4fb7-93ad-be9dc3dcd41d

Access and use of this website and the material on it are subject to the Terms and Conditions set forth at
https://nrc-publications.canada.ca/eng/copyright
READ THESE TERMS AND CONDITIONS CAREFULLY BEFORE USING THIS WEBSITE.

L’accès à ce site Web et l’utilisation de son contenu sont assujettis aux conditions présentées dans le site
https://publications-cnrc.canada.ca/fra/droits
LISEZ CES CONDITIONS ATTENTIVEMENT AVANT D’UTILISER CE SITE WEB.

Questions? Contact the NRC Publications Archive team at
PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca. If you wish to email the authors directly, please see the first page of the publication for their contact information.

Vous avez des questions? Nous pouvons vous aider. Pour communiquer directement avec un auteur, consultez la première page de la revue dans laquelle son article a été publié afin de trouver ses coordonnées. Si vous n’arrivez pas à les repérer, communiquez avec nous à PublicationsArchive-ArchivesPublications@nrc-cnrc.gc.ca.

National Research Council Canada
Conseil national de recherches Canada

A comparison of future realizations of the kilogram

M. Stock[1], P. Barat[1], P. Pinot[2], F. Beaudoux[3], P. Espel[3], F. Piquemal[3], M. Thomas[3], D. Ziane[3], P. Abbott[4], D. Haddad[4], Z. Kubarych[4], J.R. Pratt[4], S. Schlamminger[4], K. Fujii[5], K. Fujita[5], N. Kuramoto[5], S. Mizushima[5], L. Zhang[5], S. Davidson[6], R.G. Green[7], J. Liard[7], C. Sanchez[7], B. Wood[7], H. Bettin[8], M. Borys[8], I. Busch[8], M. Hämpke[8], M. Krumrey[8], A. Nicolaus[8]

[1] Bureau International des Poids et Mesures (BIPM), 92310 Sèvres, France
[2] Conservatoire National des Arts et Métiers (CNAM), 93210 La Plaine St-Denis, France
[3] Laboratoire National de Métrologie et d'Essais (LNE), 78197 Trappes, France
[4] National Institute of Standards and Technology (NIST), Gaithersburg, MD 20899, USA
[5] National Metrology Institute of Japan (NMIJ), Tsukuba, Ibaraki 305-8563, Japan
[6] National Physical Laboratory (NPL), Teddington, Middlesex, TW11 0LW, UK
[7] National Research Council of Canada (NRC), Ottawa, ON K1A 0R6, Canada
[8] Physikalisch-Technische Bundesanstalt (PTB), 38116 Braunschweig, Germany

Abstract

The definition of the kilogram in the International System of Units (SI) is expected to be revised in 2018. The present definition of the kilogram, the mass of the International Prototype of the Kilogram (IPK), adopted in 1889, would then be replaced by a definition based on a fixed numerical value of the Planck constant. The Consultative Committee for Mass and Related Quantities (CCM) has requested that, as one of the essential steps before the redefinition, a comparison of kilogram realizations based on future realization methods, Kibble[1] balances and X-ray crystal density (XRCD) experiments, be organized.

This comparison was carried out during 2016 in the form of a "Pilot Study". One aim of the study was to determine the uniformity of mass dissemination after the redefinition by comparing mass calibrations based on different future realization experiments. Another aim was to test the continuity of the mass unit across the redefinition by comparing mass calibrations based on Kibble balances and XRCD experiments with those based on the IPK. This paper describes the organization of the comparison and presents its results.

1. Introduction

It is planned that in the near future four of the seven base units of the International System of Units (SI), the kilogram, the ampere, the kelvin and the mole will be redefined [1]. It is expected that the revised SI will be approved at the meeting of the General Conference on Weights and Measures (CGPM) in November 2018 and that it will come into force on World Metrology Day 20 May 2019. The present

---

[1] The Consultative Committee for Units (CCU) decided in its meeting in June 2016 to refer to the watt balance as the "Kibble balance" in homage to Bryan Kibble, who originally conceived the idea of this experiment.

artefact definition of the kilogram, being the mass of the International Prototype of the Kilogram (IPK), would then be replaced by a definition based on a fixed numerical value of the Planck constant.

At present, experimental access to the definition of the kilogram is only possible at the International Bureau of Weights and Measures (BIPM), where the IPK is maintained. National Metrology Institutes (NMIs) of Member States of the BIPM possess national prototypes of the kilogram which are sent periodically to the BIPM for calibration traceable to the IPK. This system guarantees a globally consistent dissemination of the mass unit, since all mass calibrations are ultimately traceable to the IPK [2]. However, the definition of the kilogram relies on a material artefact, the mass stability of which cannot be guaranteed over time. Indeed, it has been observed that the mass differences between the IPK and its six official copies, which are stored with it, have changed since the first mass comparisons were made in 1889 by several tens of micrograms [2]. The new definition of the kilogram will be based on a fundamental constant of physics and will therefore be inherently stable.

The future definition of the kilogram does not imply any particular experiment for the practical realization of the mass unit. Any method capable of deriving a mass value traceable to the numerical value of the Planck constant (and to the definitions of the metre and the second) will have the potential to be a primary method. This is a method having the highest metrological properties, whose operation can be fully described and understood, for which a full uncertainty statement can be written down in terms of SI units and which does not require a reference standard of the same quantity. The Consultative Committee for Mass and Related Quantities (CCM) is responsible for publishing the *mise en pratique* of the definition of the kilogram [3], which describes the currently available primary methods for the practical realization of the kilogram. At the time of writing, it has been demonstrated that two primary methods have the capability of realizing the kilogram according to its future definition with relative standard uncertainties of a few parts in 10^8: the Kibble balance [4] and the X-ray crystal density (XRCD) method [5]. An artefact whose mass has been calibrated by a primary method becomes a primary mass standard and can be used to disseminate the mass unit to secondary mass standards by conventional weighing techniques.

The new definition will, in principle, allow any NMI to realize the kilogram. This makes the new definition universal but raises the question of the uniformity of the independent realizations. Another important aspect is the continuity between realizations based on the present definition and those based on the future definition. The kilogram should not change noticeably as a consequence of the redefinition.

The CCM has developed jointly with the CCU a *Roadmap towards the redefinition of the kilogram* [6], which identifies the main steps towards the ratification of the redefinition. One important task was the extraordinary calibration campaign, carried out at the BIPM in 2014, to provide improved traceability to the IPK for those NMIs which are involved in accurate determinations of the Planck constant [7]. This aimed to ensure that the value of the Planck constant, which will ultimately be used to redefine the kilogram, will be precisely measured in terms of the present artifact definition. Another important task on the roadmap is the Pilot Study comparing kilogram realizations using future realization methods, which is the topic of this article. A validation of the procedures for the future realization and dissemination of the kilogram is one of the conditions which the CCM formulated at its meeting in 2013

[8]. The Pilot Study is also mentioned in the draft *mise en pratique* of the (revised) definition of the kilogram [3]. Pilot studies are a category of comparison undertaken in a new field of metrology or using new techniques which are not yet considered as fully mature. They are considered as trial comparisons.

The aims of the Pilot Study were to test the uniformity of kilogram realizations based on independent realization experiments and to investigate the continuity between the kilogram realizations based on the present and the future definitions.

2. Organization of the Pilot Study

The Pilot Study was organized by the BIPM, supported by Horst Bettin (PTB, Germany), Stuart Davidson (NPL, UK) and Chris Sutton (Measurement Standards Laboratory, MSL, New Zealand). All NMIs working on realization experiments were invited to participate, under the condition that they would be able to realize the mass unit at the level of 1 kg with a standard uncertainty not larger than 200 µg, that is 2 parts in 10^7. Five institutes participated: the LNE (France), the NIST (USA), the NMIJ (Japan), the NRC (Canada) and the PTB (Germany). The LNE , NIST and NRC used Kibble balances, the NMIJ and the PTB used the ^{28}Si-spheres AVO28-S5c and AVO28-S8c from the International Avogadro Coordination [9] for their calibrations. The target calibration uncertainties initially reported by the participants ranged from 20 µg to 200 µg.

Each participant was asked to provide two independent sets of 1 kg travelling standards. One set (Set 1) consisted of one Pt-Ir standard and a second optional standard of the participant’s choice. These standards had to be calibrated as directly as possible using the realization experiment. To achieve the lowest possible uncertainty, these experiments are typically operated under vacuum, and therefore the standards of Set 1 were to be calibrated under vacuum by the participants. One possibility would be a direct calibration in a Kibble balance or in a vacuum mass comparator against an Avogadro sphere. Another possibility would be the use of an intermediate transfer standard, which would be calibrated using the realization experiment and then used in a second step to calibrate the travelling standards in a vacuum mass comparator. The mass values of the standards were to be calculated by all participants from the same value of the Planck constant. The comparison protocol stated that the value recommended by the 2014 CODATA adjustment of fundamental constants, $h = 6.626\ 070\ 040 \times 10^{-34}$ J s [10], should be used. The choice of this particular value was not relevant for the purposes of the comparison, but all NMIs needed to use the same value. However, the choice of the value will impact on the continuity between the present and the future primary realizations. It is for this reason that the CODATA value has been selected. It is important to note that the value which will ultimately be determined by the CODATA Task Group on Fundamental Constants for the revised definition of the kilogram could be different from this 2014 value.

The second set of travelling standards (Set 2) consisted of two 1 kg stainless steel standards which were to be calibrated in air, traceable to the realization experiment. This required transferring a mass standard from vacuum to air, by applying a correction for the surface sorption, and a buoyancy correction for the weighing in air, in the case that standards of different density were used.

Using these two sets of standards both the realization of the mass unit and its subsequent dissemination were evaluated by the Pilot Study. It was the objective of the Pilot Study to compare the mass units as realized and disseminated in practice by the participating institutes, using Kibble balance or XRCD experiments. The technical protocol did not therefore prescribe particular measurement procedures, but the participants were to use those calibration procedures for both sets of standards, which they planned to apply after the redefinition.

After a calibration with the realization experiments at the participating NMIs, the standards were shipped to the BIPM. Although the standards of Set 1 had been calibrated under vacuum (except for one standard sent by the LNE), they were shipped in containers filled with air, since no practical solution for shipping under vacuum exists. Mass comparisons have been carried out at the BIPM among all standards of Set 1 under vacuum (except for the standard sent by the LNE, which was linked with the others by the use of BIPM air-vacuum transfer standards) and among all standards of Set 2 in air. The results of these weighings together with the mass values and uncertainties provided by the participants allow assessment of the capability of future realization methods to realize and to disseminate the same unit of mass (see Section 4). The travelling standards of both sets were also compared with BIPM working standards, the masses of which are traceable to the IPK from comparisons made in 2014. For the weighings under vacuum of the standards of Set 1 a set of two Pt-Ir sorption artefacts, a cylinder and a stack of disks, was used to link with the BIPM working standards, kept in air. This allowed the verification of the agreement between the mass values traceable to the realization experiments and to the IPK.

As a means to verify the stability of the standards of Set 1 during transportation, the participants also determined the mass of these standards in air, traceable to their national prototype (and ultimately to the IPK), before sending the standards to the BIPM and after receiving them back. The BIPM had determined the mass of the standards in air traceable to the IPK on reception and before sending them back. Since the participants' and the BIPM's measurements were all be traceable to the IPK, this should allow detection of any significant mass changes during transportation. For the stainless steel standards of Set 2, the participants determined the mass changes during the study from comparisons with stable reference standards before and after sending them to the BIPM.

The participants carried out their calibration measurements from January to March 2016. The comparison measurements at the BIPM were made from May to July 2016, and the final stability checks at the NMIs from August to November 2016. The report on the Pilot Study, which presents all experimental data and a more detailed description of the work, was published in June 2017 [11].

3. Measurements by the participants and the BIPM

3.1 Measurements at the participating NMIs

Tables 1 and 2 show the travelling standards which were calibrated by the participants and sent to the BIPM and the standard uncertainties of their mass values.

Table 1: Travelling standards of Set 1 and standard uncertainty of their mass value under vacuum (for n° 13 in air).

Institute	**Identification of standard**	**Type and manufacturer**	**Standard uncertainty of calibration / mg**
LNE	n° 13	Pt-Ir prototype (BIPM)	0.140 (in air)
NIST	K104	Pt-Ir prototype (BIPM)	0.0359
	141714	stainless steel (Mettler Toledo)	0.0279
NMIJ	n° 94	Pt-Ir prototype (BIPM)	0.0238
	E59	Pt-Ir mass standard (Stanton Instrum.)	0.0238
NRC	K50	Pt-Ir prototype (BIPM)	0.015
PTB	Pt109	Pt-Ir prototype (BIPM)	0.019
	Si14-02	natural Si-sphere (PTB)	0.019

Table 2: Stainless steel travelling standards of Set 2 and standard uncertainty of their mass value in air.

Institute	**Identification of standard**	**Type and manufacturer**	**Standard uncertainty of calibration / mg**
LNE	E	OIML E0 (Mettler Toledo)	0.140
	INM	OIML E2 (Mettler Toledo)	0.140
NIST	Zwiebel 7	OIML E1 with lifting knob (Zwiebel)	0.0368
	Zwiebel 8	OIML E1 with lifting knob (Zwiebel)	0.0368
NMIJ	S1_2	right cylinder (Chyo balance)	0.0255
	S2_1	right cylinder (Chyo balance)	0.0255
NRC	HSA2	right cylinder (Häfner)	0.015
	HSA3	right cylinder (Häfner)	0.015
PTB	D1	cylinder (PTB)	0.019
	D2	cylinder (PTB)	0.019

The following sections describe briefly how the mass calibrations were carried out by the participants.

LNE: At the time when the measurements for the Pilot Study were made, the LNE Kibble balance was not yet operating under vacuum. In addition, the prototype n° 13 showed a very regular linear drift with time over more than a century and should not be used under vacuum.

A 500 g Pt-Ir mass standard was calibrated in air on the LNE Kibble balance. It served for the calibration of a similar 500 g Pt-Ir mass standard using an M_one mass comparator. Both standards were then used together to determine the mass in air of the travelling standard n° 13.

Both standards were used together to calibrate a 1 kg Pt-Ir standard, which served to determine the mass of a stainless steel standard. This in turn was used to calibrate the two stainless steel standards of Set 2. The buoyancy correction was made using artefacts.

NIST: The travelling standards of Set 1 were measured under vacuum directly in the NIST-4 Kibble balance. The air-to-vacuum correction of K104 was determined using a set of sorption artefacts, which have the same surface characteristics as K104. The travelling standards of Set 2 were compared directly with K104 in air. The buoyancy correction was made using the CIPM-2007 formula [12].

NMIJ: The NMIJ used the sphere AVO28-S5c from the International Avogadro Coordination. A new determination of the core volume was carried out by optical interferometry. The mass of the surface layers was determined by new XPS and ellipsometry investigations. For the lattice constant, the relative atomic mass, and the influence of point defects, the results from the previous determination of the Avogadro constant were used [9]. The travelling standards of Set 1 were compared with AVO28-S5c in a vacuum mass comparator.

For the calibration of the standards of Set 2, the mass of AVO28-S5c in air was determined using sorption artefacts. The sphere was then used to calibrate a stainless steel standard. The buoyancy correction was carried out using a set of buoyancy artefacts. The stainless steel standard served to calibrate the two travelling standards S1_2 and S2_1. For these measurements, the buoyancy correction was made with the CIPM-2007 formula.

NRC: Vacuum cycling experiments were performed to both stabilize and determine the stability of six masses during air-vacuum cycling. Masses were found to be stable to within ± 0.4 µg after three cycles. One of the masses was then extracted from the vacuum balance and transferred in air to be calibrated under vacuum in the Kibble balance. After calibration, the transfer standard was reinserted into the vacuum balance via the load-lock and compared again with the five other standards of the first measurement series which had remained continuously under vacuum. The transfer mass showed an average increase of 8.8 µg with respect to the five masses and therefore a correction of – 4.4 µg (with the same magnitude uncertainty) was applied to the calibrated value. During this series, the transfer standard was used to calibrate two 500 g standards, both made of stainless steel. Finally, both 500 g standards were stacked together to calibrate the 1 kg travelling standard K50 under vacuum. All comparisons were performed in an M_one mass comparator.

For the calibration of the standards of Set 2, a set of stainless steel sorption artefacts were compared with each other under vacuum as well as to K50 and the stack traceable to the realization experiment. The stainless steel sorption artefacts were vented to air, corrected for their sorption, and used to calibrate the stainless steel travelling standards in air. The buoyancy correction was carried out using the CIPM-2007 formula to determine the air density.

PTB: The PTB used the sphere AVO28-S8c from the International Avogadro Coordination. Only the volume of the sphere and the surface layers were measured anew. For the other parameters, the results from the previous determination of the Avogadro constant were used [9]. The travelling standards of Set 1 were compared with AVO28-S8c under vacuum in an M_one mass comparator.

A Pt-Ir cylinder was calibrated against AVO28-S8c in vacuum. This cylinder is one of the two standards used as sorption artefacts for the link between the mass of AVO28-S8c under vacuum and the mass of the two travelling standards of Set 2 in air. The buoyancy correction was made using the artefact method.

Since the PTB used the sphere AVO28-S8c and the NMIJ the sphere AVO28-S5c, their results are partly correlated. The core volumes and surface layers have been independently re-determined in both institutes. The point defect corrections are different and only partly correlated. All other correlated uncertainty contributions have been estimated by the PTB as 8 µg and are considered to be the same as in the measurements of NMIJ, resulting in a correlation coefficient of about 13 % for the masses of the spheres. It has been verified that this correlation is negligible for the calculation of the reference value of this comparison (Section 4).

3.2 Comparison measurements at the BIPM

Upon arrival all standards of Set 1 were compared in air with two BIPM Pt-Ir working standards, which were traceable to the IPK. A set of artefacts was used to determine air density to allow the calculation of buoyancy corrections. These results were compared with the participants' calibrations against their national prototypes, to allow an estimation of possible mass changes during transportation. The BIPM weighings were repeated before the standards were sent back to the participants.

The travelling standards of Set 1 (except n° 13 of the LNE) were then compared with each other, including a set of BIPM Pt-Ir sorption artefacts, under vacuum in an M_one mass comparator. The sorption artefacts were calibrated traceable to the IPK in air. In total 120 mass differences were determined and a least-squares adjustment was performed to assign masses to the seven travelling standards of Set 1. Prototype n° 13 of the LNE was weighed in air and compared with the other standards via the sorption artefacts, which were used in air and under vacuum. As requested by the PTB, the Si-sphere was washed three times by applying the method developed by the NMIA (National Measurement Institute, Australia) [13] before the vacuum measurements. All other standards were only gently brushed to remove dust.

The stainless steel standards of Set 2 were compared with two BIPM stainless steel standards, which were traceable to the IPK. A set of 150 mass differences served as input data for the determination of the masses of the 10 travelling standards.

Before and after the comparison measurements, the mass stability of the BIPM working standards used for these weighings was verified by comparison with other working standards, which had not been used during this period. In general the mass stability was better than 0.002 mg and corrections from linear interpolations were applied.

4. Results of the Pilot Study

4.1 Results of the comparison of calibrations under vacuum (Set 1)

As a first step in determining the consistency of the NMIs' calibrations, the masses of the travelling standards of Set 1 under vacuum were determined with reference to the BIPM working standards, kept in air and calibrated in terms of the BIPM as-maintained mass unit. This was achieved by using a set of two Pt-Ir sorption standards, a cylinder and a stack of disks, which was included in each vacuum weighing set. Both standards had also been compared with the BIPM working standards in air. In a second step the differences between the mass values attributed by the NMIs (based on their realization experiments) and the BIPM (traceable to the IPK) were calculated:

$$\Delta m_{i,j} = m_{i,j}^{\mathrm{NMI}} - m_{i,j}^{\mathrm{BIPM}} \qquad (1)$$

where the index *i* stands for the laboratory, and *j* for one of its travelling standards. In this way the as-maintained BIPM mass unit served as a common reference to compare the NMIs' calibrations. The uncertainty of $\Delta m_{i,j}$ includes the calibration uncertainty reported by the laboratory (Table 1), the uncertainty of the comparison at the BIPM (which is negligible for the purposes of this comparison) and two contributions characterizing potential mass changes of the travelling standards during transportation and due to the repeated vacuum-air cycling. The standard uncertainty related to transport was estimated for each standard as:

$$u_{\mathrm{transport}} = \frac{1}{2}\left(\left|m_{\mathrm{BIPM}}^{\mathrm{arrival}} - m_{\mathrm{NMI}}^{\mathrm{initial}}\right| + \left|m_{\mathrm{NMI}}^{\mathrm{final}} - m_{\mathrm{BIPM}}^{\mathrm{departure}}\right|\right) \qquad (2)$$

Where $m_{\mathrm{BIPM}}^{\mathrm{arrival}}$ and $m_{\mathrm{BIPM}}^{\mathrm{departure}}$ are the masses of the standard determined at the BIPM at arrival and departure; $m_{\mathrm{NMI}}^{\mathrm{initial}}$ and $m_{\mathrm{NMI}}^{\mathrm{final}}$ are the mass values attributed by the NMI before and after sending the standard to the BIPM. These masses are all traceable to the IPK, and not to the realization experiments.

For all travelling standards, except for 141714 sent by the NIST, equation (2) leads to very similar uncertainties, on average 0.004 mg with a standard deviation of 0.001 mg. Therefore a common transport uncertainty of 0.004 mg was used for all standards, except for 141714, where the use of equation (2) gives an uncertainty of 0.0214 mg. This stainless steel standard had already shown some instability during the calibrations at the NIST and again during its stay at the BIPM.

The uncertainty of the Pilot Study is also influenced by the mass stability of the travelling standards under repeated air-vacuum transfers which were inevitable because the standards had to be shipped in air. The relevant quantity is the change of the vacuum mass between the NMI and the BIPM. The data from the Pilot Study does not allow assessment of this quantity directly. It has therefore been agreed by the participants to estimate this uncertainty based on previous experience: NRC stated in its measurement report a cycling/short term stability of 0.0025 mg. The BIPM made the experience that after a small number of cycles the vacuum mass becomes stable to within 0.001 mg. For the purposes of the Pilot Study an uncertainty of 0.002 mg has been estimated for Pt-Ir standards. Since the stainless steel standard 141714 has about twice the surface of a Pt-Ir standard, an uncertainty of 0.004 mg was assumed. For the Si sphere, the PTB reported a mass stability of 0.005 mg, including the effect of repeated cleaning. Both the transport uncertainty and the air-vacuum cycling uncertainty are small compared to the uncertainty of the realization experiments (Table 1).

For NMIs which sent two standards, the mean of both differences $\Delta m_{i,1}$ and $\Delta m_{i,2}$ and its uncertainty were calculated, by taking into account the correlations. In this way, for each participant one mass difference Δm_i and its uncertainty were obtained. The comparison reference value was calculated as the weighted mean of these differences. Finally, the deviations of the NMIs' results from the reference value, $\Delta m_i'$, were determined. Since the BIPM calibrated the travelling standards traceable to the IPK, the deviation between the BIPM calibrations and the weighted mean of the five participants' calibrations can also be calculated. The results are shown in Figure 1 and Table 3.

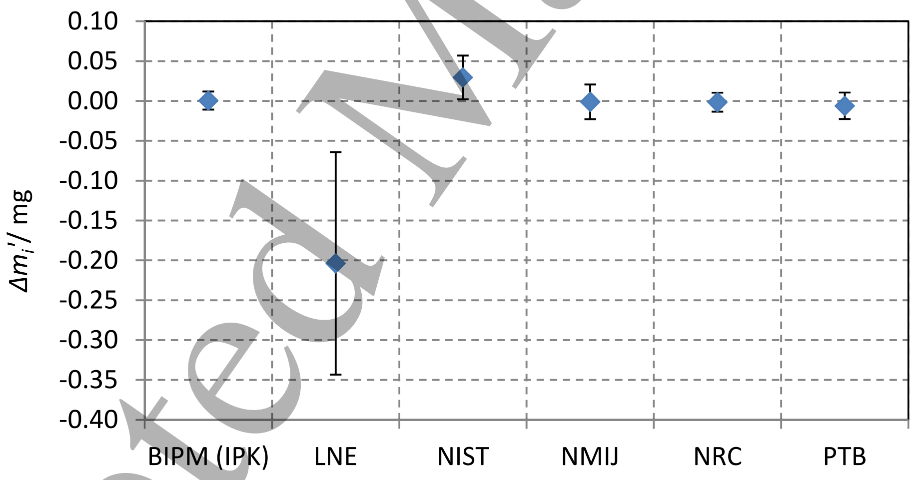

Fig 1: Deviations of the NMIs' results for the calibration of a 1 kg standard under vacuum, using the realization experiment, from the reference value and related standard uncertainties. The reference value is calculated as the weighted mean of the five NMIs' results. The standard uncertainty of the reference value is 0.010 mg. The difference between the BIPM calibration based on the IPK and the reference value is also indicated.

Table 3: Deviations of the NMIs' results for the calibration of a 1 kg standard under vacuum from the reference value and related standard uncertainties. The difference between the BIPM calibration based on the IPK and the reference value is also indicated.

Institute	**Deviation from reference value** $\Delta m_i'$ **/ mg**	**Standard uncertainty of deviation** $u(\Delta m_i')$ **/ mg**
BIPM (IPK)	0.0006	0.0113
LNE	-0.2038	0.1396
NIST	0.0296	0.0274
NMIJ	-0.0012	0.0218
NRC	-0.0015	0.0119
PTB	-0.0061	0.0165

The uncertainty of the reference value is 0.010 mg. The Birge ratio of the five participants' results is 0.90. A Birge ratio below one indicates consistent data for the realization part of the Pilot Study. This is confirmed by the fact that for four participants the uncertainty bars overlap, which is also the case for the LNE for a coverage level of $k = 2$. The calibration based on the as-maintained mass unit of the BIPM is consistent with the calibrations based on Kibble balances and Avogadro spheres.

The realization experiments of the participants were used to calibrate the travelling mass standards. However, the same measurements can be used to derive a value of the Planck constant, if the mass of the travelling standards is known in terms of the present kilogram definition. A comparison of the Planck constant values obtained during this work leads to results which are highly consistent with those shown in Fig. 1 and Table 3. The differences between the NMIs' mass calibrations and Planck constant values are consistent at the level of several parts in 10^9. This observation is interesting because it demonstrates that the outcome of the Pilot Study is not dominated by the behavior of the travelling standards, but that the observed mass differences reflect differences between the realization experiments.

4.2 Results of the comparison of calibrations in air (Set 2)

The determination of the degrees of equivalence of the NMIs' calibrations for the stainless steel standards followed the same approach as described in Section 4.1 for the standards of Set 1.

The uncertainty component related to a potential mass change during transportation was estimated by the participants, who weighed the travelling standards before sending them to the BIPM and after receiving them back against other reference standards. Interestingly, some of the stainless steel standards lost a significant amount of mass, while others were stable. The standards of NMIJ and NRC were stable within a few micrograms, whereas the PTB standards both lost 0.009 mg and the NIST standards lost 0.009 mg and 0.023 mg. One of the LNE standards showed a mass loss of 0.054 mg, which was considered as anomalous. In the data treatment the LNE calibration result obtained before the comparison was corrected for this change, because this led to a result which was consistent with that of the second LNE standard. It might be that the initial calibration result had been influenced by a density determination in a hydrostatic balance which was carried out a couple of days before. Since the observed changes are the cumulative effect of the trip to the BIPM and back to the NMI, they were taken as an upper limit and the transport uncertainty was derived by dividing their absolute value by $\sqrt{3}$.

The comparison reference value for the stainless steel standards was calculated in the same way as for the standards of Set 1. The deviations between the NMIs' calibrations and the reference value are shown in Figure 2 and in Table 4. Since the BIPM calibrated the travelling standards traceable to the IPK, the deviation between the BIPM calibrations and the weighted mean of the five participants' calibrations can also be calculated and is included in the figure.

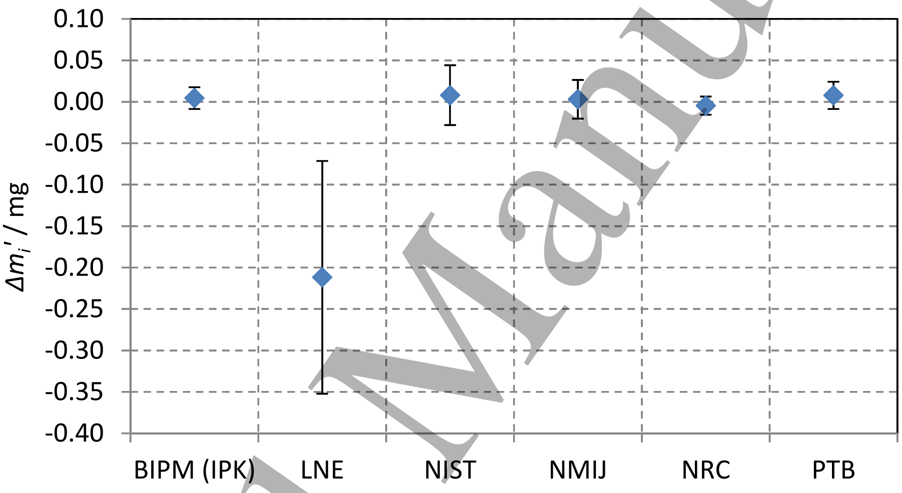

Fig. 2: Deviations of the NMIs' results for the calibration of a 1 kg stainless steel standard in air from the reference value and related standard uncertainties. The difference between the BIPM calibration based on the IPK and the reference value is also indicated. The standard uncertainty of the reference value is 0.010 mg.

The uncertainty of the reference value is 0.010 mg, identical as for the standards of Set 1. The Birge ratio of the five participants' results is 0.80, which indicates that the data are consistent.

A comparison with Fig. 1 shows that the results of the calibrations under vacuum and in air are very similar. This indicates that the vacuum-to-air transfer and the buoyancy corrections are at the present state not the limiting factors.

Table 4: Deviations of the NMIs' results for the calibration of a 1 kg standard in air from the reference value and related standard uncertainties.

Institute	Deviation from reference value $\Delta m_i'$ / mg	Standard uncertainty of deviation / mg
BIPM (IPK)	0.0045	0.0131
LNE	-0.2118	0.1405
NIST	0.0080	0.0360
NMIJ	0.0031	0.0233
NRC	-0.0046	0.0109
PTB	0.0077	0.0164

5. Summary and conclusions

The CCM Pilot Study of future realizations of the kilogram has been carried out to test the consistency of kilogram realizations based on Kibble balances and Avogadro spheres, and to test the continuity between the present and future realizations. The experience gained from the Pilot Study will be useful in developing the protocol for future key comparisons of realization experiments.

It has been found that calibrations of 1 kg mass standards under vacuum, using the future realization experiments as directly as possible agree for four out of five participants within the estimated standard uncertainties, and within the expanded uncertainty (k = 2) with the fifth. The standard uncertainty of the weighted mean of calibrations from the five experiments is 0.010 mg. The agreement between calibrations based on the IPK and the weighted mean is at the level of 0.001 mg.

The comparison of calibrations of 1 kg stainless steel standards in air shows the same level of agreement. The vacuum-to-air transfer and the buoyancy corrections are no limiting factors for the present uncertainty level of realization experiments.

The result of the comparison of mass calibrations is highly consistent with the comparison of determinations of the Planck constant, obtained during the same measurements. This indicates that at present, the results of the Pilot Study are not dominated by the behavior of the travelling standards. This issue might become more significant for future comparisons, when the uncertainty of the realization experiments will become smaller.

References

[1] Milton M.J.T., Davis R., Fletcher N., Towards a new SI: a review of progress made since 2011, *Metrologia*, 2014, **51(3)**, R21-R30.

[2] Davis R., The SI unit of mass, *Metrologia*, 2003, **40(6)**, 299-305.

[3] *Mise en pratique* for the definition of the kilogram, version 9.0, working document CCM/15-02A, http://www.bipm.org/cc/CCM/Allowed/15/02A_MeP_kg_141022_v-9.0_clean.pdf.

[4] Robinson I.A., Schlamminger S. , "The watt or Kibble balance: a technique for implementing the new SI definition of the unit of mass", *Metrologia* 2016 **53** A46-A74.

[5] Fujii K. et al., "Realization of the kilogram by the XRCD method", *Metrologia* 2016 **53** A19-A45.

[6] Joint CCM and CCU roadmap towards the redefinition of the SI in 2018, http://www.bipm.org/utils/common/pdf/SI-roadmap.pdf.

[7] Stock M., Barat P., Davis R.S., Picard A., Milton M.J.T., Calibration campaign against the international prototype of the kilogram in anticipation of the redefinition of the kilogram part I: comparison of the international prototype with its official copies, *Metrologia*, 2015, **52(2)**, 310-316.

[8] Recommendation G1 (2013), Report of the 14th meeting of the Consultative Committee for Mass and Related Quantities (CCM), http://www.bipm.org/utils/common/pdf/CC/CCM/CCM15.pdf .

[9] Azuma Y., et al., Improved measurement results for the Avogadro constant using a 28Si-enriched crystal, *Metrologia*, 2015, **52**, 360-375.

[10] Mohr P.J., Newell D.B., Taylor B.N., CODATA Recommended Values of the Fundamental Physical Constants: 2014, *Rev. Mod. Phys.*, 2016, **88**, 035009 or arXiv:1507.07956v1.

[11] Stock M. et al., Report on CCM Pilot Study CCM.R-kg-P1, Comparison of future realizations of the kilogram, 27 June 2017, http://www.bipm.org/cc/CCM/Allowed/16/03-7B2_CCM-PilotStudy-FinalReport.pdf

[12] Picard A., Davis R.S., Gläser M., Fujii K., Revised formula for the density of moist air (CIPM-2007), *Metrologia* **45** (2008) 149-155.

[13] Picard A., Barat P., Borys M., Firlus M., Mizushima S., State-of-the-art mass determination of ^{28}Si spheres for the Avogadro project, *Metrologia*, 2011, **48**, S112-119.