

## **Air-vacuum transfer; establishing traceability to the new kilogram**

**Stuart Davidson<sup>1</sup>, James Berry<sup>1</sup>, Patrick Abbott<sup>2</sup>, Kilian Marti<sup>3</sup>, Richard Green<sup>4</sup>, Andrea Malengo<sup>5</sup>, Lars Nielsen<sup>6</sup>**

<sup>1</sup> National Physical Laboratory NPL, Hampton Road, Teddington, Middlesex TW11 0LW, UK

<sup>2</sup> National Institute of Standards and Technology NIST, 100 Bureau Drive, Stop 8221, Gaithersburg, MD 20899-1070, USA

<sup>3</sup> Federal Institute of Metrology METAS, Lindenweg 50, 3003 Bern, Switzerland

<sup>4</sup> National Research Council NRC, 1200 Montreal Road, Building M-58, Ottawa, Ontario K1A 0R6, Canada

<sup>5</sup> Istituto Nazionale di Ricerca Metrologica INRIM, Strada delle Cacce, 91 - 10135 Torino, Italy

<sup>6</sup> Danish National Metrology Institute DFM, Matematiktorvet 307, 1.sal, DK-2800 Lyngby, Denmark

### **Abstract**

The redefinition of the kilogram, along with another three of the base SI units, is scheduled for 2018. The current definition of the SI unit of mass assigns a mass of exactly one kilogram to the International Prototype of the kilogram, which is maintained in air and from which the unit is disseminated. The new definition, which will be from the Planck constant, involves the realisation of the mass unit in vacuum by the watt balance or Avogadro experiments. Thus, for the effective dissemination of the mass unit from the primary realisation experiments to end users, traceability of mass standards transferred between vacuum and air needs to be established and the associated uncertainties well understood. This paper describes means of achieving the link between a unit realised in vacuum and standards used in air, and the ways in which their use can be optimised. It also investigates the likely uncertainty contribution introduced by the vacuum-air transfer process.

### **1. Background**

As discussed elsewhere in this issue, the kilogram, the last of the base SI units defined in terms of a physical artefact, will soon be replaced with a definition based on a fundamental constant of nature. As things stand the new kilogram will be defined in terms of the Planck constant and realised using the watt balance [1] or X-ray crystal density (Avogadro) experiments [2].

While the redefinition and subsequent traceability to a constant of nature in theory ensures the long-term stability of the SI unit of mass, a number of issues arise from the redefinition. One of the aims of moving away from an artefact-based definition is to remove the limitations of a single point of traceability on the realisation and worldwide dissemination of the unit. Although this will in theory be achieved, there are in fact currently only two [3,4] experiments that can achieve the level of uncertainty of  $2 \times 10^{-8}$  required for the successful redefinition and subsequent dissemination of the SI unit of mass. A further issue, which will need to be addressed for the successful implementation of the kilogram redefinition, is the necessity to establish robust traceability from the mass unit in air to mass realised in vacuum. This is

<sup>1</sup>necessary both for the effective redefinition of the unit with traceability to the International Prototype Kilogram (IPK) and for the dissemination of the unit from the primary realisation experiments following redefinition.

The unit of mass is currently defined as follows: “The kilogram is the unit of mass; it is equal to the mass of the international prototype of the kilogram.” The BIPM SI Brochure [5] gives the additional information that the definition refers to the mass of the IPK immediately after cleaning and washing and that it is kept (and used) at the BIPM under ambient conditions.

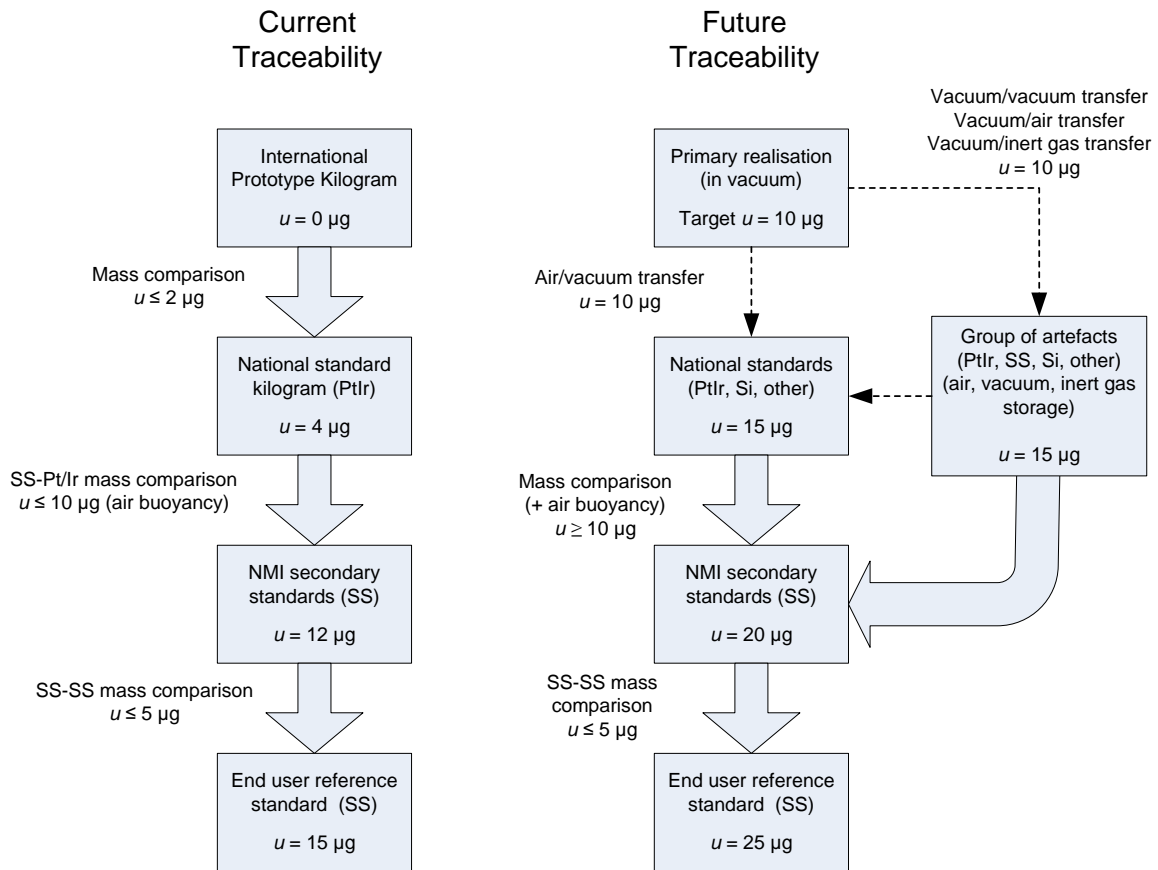
The nature of the watt balance and X-ray crystal density (XRCD) experiments means that both methods will realise the kilogram in vacuum. In order to fix the Planck and Avogadro constants with relation to the IPK, it is necessary to make as direct a link as possible between the IPK in air and the watt balance and XRCD realisations of the kilogram in vacuum. This requires an in-depth understanding of the effect of moving a mass between air and vacuum and the procedures and technical infrastructure to make the transfer repeatably and with the lowest contribution to the overall uncertainty of transferring a value between the IPK and the watt balance and XRCD experiments.

## **2. Traceability post-redefinition**

In addition to the requirement for characterised air-vacuum transfer in order to fix the Planck and Avogadro constants with relation to the IPK, the post-redefinition dissemination of the kilogram, with using primary methods<sup>1</sup>, will require the mass unit to be transferred from vacuum to air. Figure 1 shows the pre- and post-redefinition traceability routes for the SI unit of mass and illustrates the added complexity, and therefore increased uncertainty, of dissemination from the new kilogram definition. The potential increase in the (uncertainties of the) Calibration and Measurement Capability (CMC) of National Measurement Institutes, and therefore of the uncertainty provided to primary end users, is significant and must be minimised by ensuring that the additional contribution due to the uncertainty of the vacuum to air transfer is as small as possible.

---

<sup>1</sup> See Section 2 of the draft *mise en pratique* of the definition of the kilogram [http://www.bipm.org/cc/CCM/Allowed/15/02A\\_MeP\\_kg\\_141022\\_v-9.0\\_clean.pdf](http://www.bipm.org/cc/CCM/Allowed/15/02A_MeP_kg_141022_v-9.0_clean.pdf) for details of the primary methods to realize the definition of the kilogram

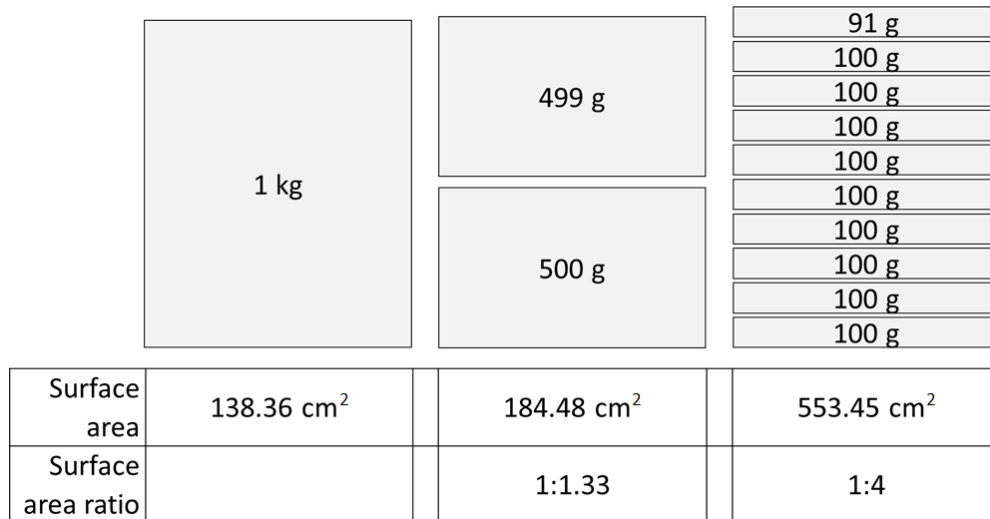


**Figure 1. Current and possible future traceability routes for the dissemination of the kilogram from the primary realisations. All uncertainties are given at the  $k = 1$  level. (PtIr : platinum-iridium, SS : stainless steel, Si : silicon, W : tungsten)**

### 3. Air vacuum transfer methods

#### 3.1 Sorption artefacts

The effect of transfer between air and vacuum on mass standards has traditionally been characterised using a set of two or more “sorption artefacts” [6]. The artefacts used should be matched in all parameters (mass, volume, surface finish) but have different geometric surface areas. This is generally achieved by using one integral artefact, usually a cylindrical kilogram standard, and further composite artefacts assembled from a series of discs. Figure 2 illustrates the typical features of a set of sorption artefacts. In order to expose all surfaces to ambient conditions the stacks of discs need to be separated. In practice this is usually achieved with short lengths of 1 mm diameter wire of nominally the same material as the discs themselves.



**Figure 2. Example of design and specifications for a set of sorption artefacts**

By comparing the mass of such sorption artefacts in vacuum and in air the measured difference in mass can be used to calculate the surface sorption per unit surface area based on the difference in the surface areas of the artefacts.

$$P_S = \frac{(m_{\text{stk(vac)}} - m_{\text{int(vac)}}) - (m_{\text{stk(air)}} - m_{\text{int(air)}})}{A_{\text{stk}} - A_{\text{int}}} \dots \dots \dots (1)$$

Where:

|                       |  |
|-----------------------|--|
| $P_S$                 | is the surface sorption per unit surface area  |
| $m_{\text{int(vac)}}$ | is the mass of the integral artefact in vacuum |
| $m_{\text{stk(vac)}}$ | is the mass of the stack of discs in vacuum    |
| $m_{\text{int(air)}}$ | is the mass of the integral artefact in air    |
| $m_{\text{stk(air)}}$ | is the mass of the stack of discs in air       |
| $A_{\text{int}}$      | is the surface area of the integral weight     |
| $A_{\text{stk}}$      | is the surface area of the stack of discs      |

The calculation of surface sorption assumes that the sorption per unit surface area is the same for the two sorption artefacts. In practice the absolute mass values of the individual artefacts in vacuum and in air do not need to be known, it is only the determined difference in mass between the two artefacts in the respective media,  $\Delta m_{\text{vac}}$  and  $\Delta m_{\text{air}}$ , which is important. Thus we can use;

$$P_S = \frac{\Delta m_{\text{vac}} - \Delta m_{\text{air}}}{A_{\text{stk}} - A_{\text{int}}} \dots \dots \dots (2)$$

The long-term stability of the sorption artefacts is therefore not critical in the determination of their surface sorption at any given time. The mass comparisons can be made with an uncertainty of better than 1  $\mu\text{g}$  and since the volumes are nominally equal the uncertainty due to the determination of air density is at the time of weighing in air negligible. In practice the major sources of uncertainty are the absolute volumes of the artefacts, or more accurately the uncertainty in the volume difference between the artefacts. Figure 3 shows an integral and 4 piece pair of platinum-iridium sorption artefacts.



**Figure 3. Integral and four piece sorption artefacts manufactured from platinum-iridium**

### **3.2 Balance incorporating magnetic coupling between air and vacuum**

An alternative to estimating the effect of air-vacuum transfer gravimetrically using sorption artefacts would be to make a direct gravimetric comparison of a mass in air with a mass in vacuum. The most obvious way to achieve this is by using magnetic coupling between a mass comparator in air and a balance pan, onto which masses can be loaded, in vacuum.

Magnetic levitation has been used in mass and density metrology for a number of years. The principle has been used in two main areas. The first application is for (fluid) density determination where a balance in air has been magnetically coupled to a sinker in a liquid [7, 8]. The second application has been used for the evaluation of the change in mass of samples in controlled conditions (pressure, vacuum, gas) [9]. While the resolution of such apparatus can be as good as 0.1  $\mu\text{g}$  the loads used have been a few tens of grams at most. In order to achieve the accuracy of 1 in  $10^8$  or better necessary to successfully disseminate the kilogram from a realisation in vacuum, a relative improvement in the uncertainty of at least an order of magnitude is necessary. To this end NIST are developing a Magnetic Suspension Mass Comparator to operate at the levels of accuracy necessary to disseminate the new kilogram, details are given in Section 5 of this paper.

#### 4. Previous gravimetric determination of surface sorption

Previous work on evaluating sorption effects on materials commonly used for primary mass standards, i.e. platinum-iridium, silicon and stainless steel, has been performed by Schwartz [10], Picard and Fang [11], Davidson [12] and Berry et al [13]. Recently published work by Berry and Davidson [14] has also examined the correlation between the measured sorption coefficients and pressure for platinum-iridium, stainless steel, and silicon artefacts. These published sorption values for the different materials and the level of vacuum at which the measurements were performed are summarised in table 1 with additional information on the time since the artefacts were cleaned. It can be seen that there is a significant variation between the sorption values reported.

NRC has performed sorption measurements on platinum-iridium prototype K50 using the NPL four piece sorption artefact shown in figure 3. K50 was subsequently sent to BIPM for a second determination of its sorption using the BIPM 10 disk stack sorption artefacts. The results of four air-vacuum cycles at NRC and BIPM are shown in figure 4. They demonstrate good agreement with an average deviation between measurements of less than  $0.005 \mu\text{g cm}^{-2}$ , well below the measurement uncertainty.

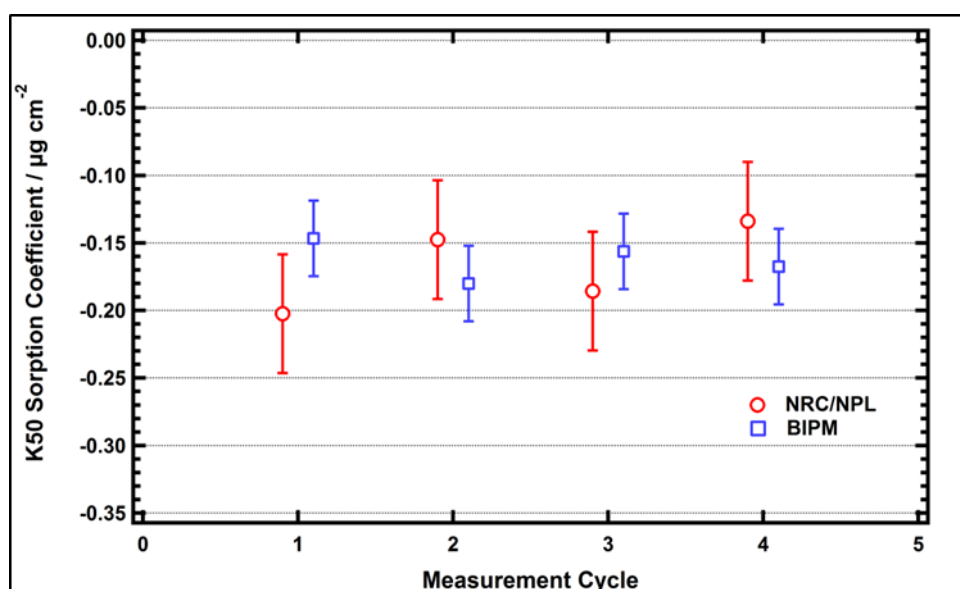


Figure 4. Sorption of NRC K50 vs. NPL and BIPM sorption stacks.

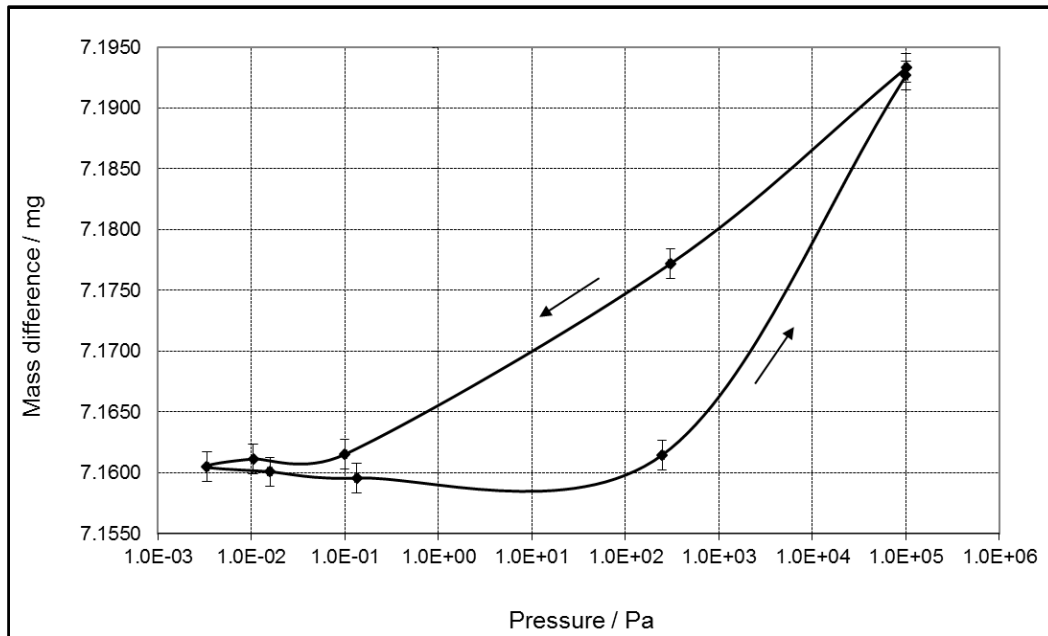
**Table 1.** Published sorption values of the materials used and pressure level of the measurements.

| Author(s)               | Material                    | Surface Roughness (R <sub>z</sub> ) / nm | Sorption value / $\mu\text{g cm}^{-2}$ | Pressure / Pa            | Time between cleaning and measurement |
|-------------------------|-----------------------------|--|--|--------------------------|---------------------------------------|
| Schwartz [9]            | stainless steel             | 120                                      | -0.030 <sup>1</sup>                    | 0.1                      | < 8 months                            |
| Schwartz [9]            | stainless steel (uncleaned) | 120                                      | -0.076 <sup>1</sup>                    | 0.1                      | -                                     |
| Picard & Fang [10]      | Pt-Ir                       | 10-100                                   | -0.080                                 | 0.1                      | < 1 month                             |
| Picard & Fang [10]      | silicon                     | <10                                      | -0.030                                 | 0.1                      | < 1 month                             |
| Picard & Fang [10]      | stainless steel             | 10                                       | -0.040                                 | 0.1                      | < 1 month                             |
| Davidson [11]           | Pt-Ir                       | 10                                       | -0.162                                 | 10 <sup>-4</sup>         | < 6 months                            |
| Davidson [11]           | stainless steel             | 60                                       | -0.154                                 | 10 <sup>-4</sup>         | < 6 months                            |
| Berry <i>et al</i> [12] | stainless steel             | 60                                       | -0.13 to -0.25                         | 0.05 to 10 <sup>-4</sup> | < 3 years                             |
| Berry & Davidson [13]   | Pt-Ir                       | 10                                       | -0.070                                 | 0.1 to 10 <sup>-3</sup>  | < 5 years                             |
| Berry & Davidson [13]   | silicon                     | <10                                      | -0.050                                 | 0.1 to 10 <sup>-3</sup>  | < 5 years                             |
| Sanchez et al [3]       | Pt-Ir                       | 10-100                                   | -0.162                                 | 0.1 to 10 <sup>-3</sup>  | > 20 years                            |

<sup>1</sup>Corrected to 50 % RH using equations (1) and (2) in [15]

The published data show a significant variation between the sorption values reported. The results reported by Schwartz indicate that there is a correlation between surface cleanliness and sorption, with “uncleaned” artefacts giving a sorption value about 2.5 times that of cleaned artefacts. There is also general correlation between the level of vacuum at which the desorption is measured and the sorption value reported.

The work of Schwartz measured surface sorption over a range of pressure from atmospheric to 0.1 Pa and found evidence of hysteresis in the sorption over this pressure range. The subsequent study by Berry and Davidson confirmed hysteretic sorption over this pressure range but additionally reported that at lower pressures, 0.1 Pa to 0.001 Pa, no significant additional desorption/sorption was observed. Figure 5 illustrates the typical reversible sorption isotherm for stainless steel as measured by Berry and Davidson. The small difference between the desorption and re-adsorption legs of the isotherm at pressures below 0.1 Pa is attributed to a (very small) time dependent desorption effect by which the surface in question desorbs more strongly bonded molecules with increasing time in vacuum. Such a time-dependent desorption phenomenon was also noted by Schwartz for pressures below 0.1 Pa.



**Figure 5. Sorption isotherm for stainless steel determined by Berry and Davidson**

## **5. NIST Magnetic Suspension Mass Comparator for vacuum-to-air mass transfer**

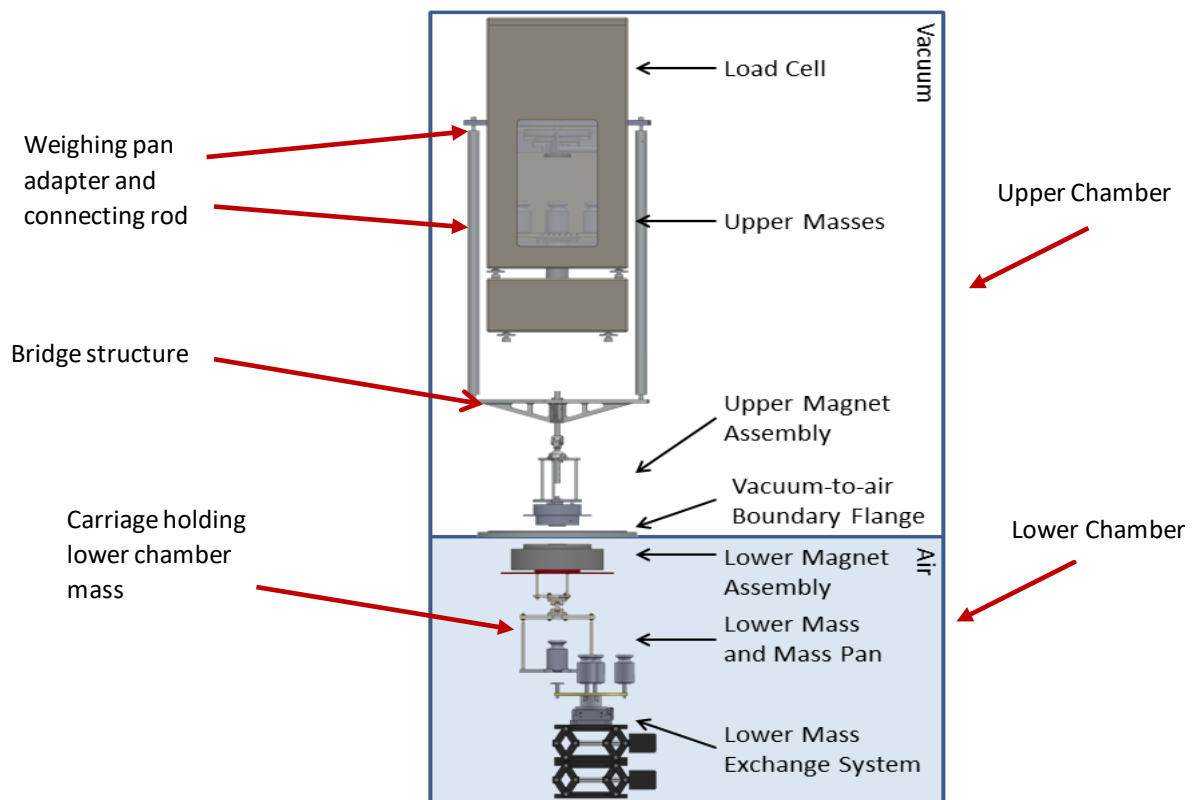
The NIST Magnetic Suspension Mass Comparator (MSMC) was designed and constructed to provide the ability to perform a direct comparison of a mass artefact in vacuum to a mass artefact in air using a single high resolution mass comparator. The MSMC will be used to transfer the realization of the kilogram in vacuum from the NIST-4 [16] watt balance to working standards in air that are used for the dissemination of mass to the United States Measurement System. Unlike the method of sorption artefacts, which determines a correction factor by measuring the amount of water that is desorbed from special masses in going from air to vacuum, the MSMC does not depend on any empirical models, nor does it require the artefacts being compared to be made of the same material or have the same surface finish. In addition, the MSMC was built with the flexibility of being able to make mass comparisons between vacuum and air, air and air, or vacuum and vacuum.

In order to use the same comparator with mass artefacts in two different environments, it is necessary to be able to independently connect each of the artefacts to the comparator's weighing pan. Since no solid means of attachment can span the vacuum-air boundary, a controlled coupling magnetic force, or magnetic suspension, was chosen; this is a feasible solution and has been made use of in the past in weighing and precision gravimetry applications [17, 18, 19, 20, 21]. Magnetic suspension of ferromagnetic objects requires a time-varying magnetic field. It was proven in the early 19th century by Earnshaw [22] that no stationary object made of charges, magnets and masses in a fixed configuration can be held in stable equilibrium by any combination of static electric, magnetic, or gravitational forces (that is, by any forces derivable from a potential satisfying Laplace's equation). A consequence of this is that stable equilibrium of any object requires its energy to possess a minimum [23].

The MSMC uses both permanent magnets (dc field) and a time-dependent magnetic field provided by a solenoid. The solenoid is part of a feedback loop that includes a Hall sensor that monitors the suspending magnetic field and corrects it to a pre-determined value to maintain suspension at a given vertical location.

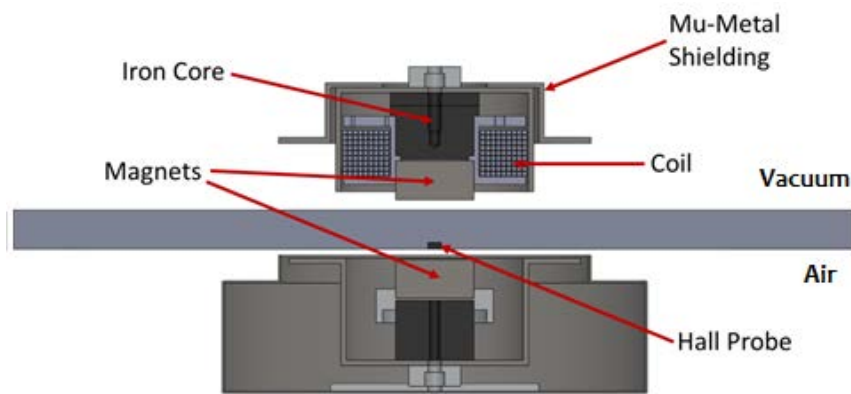
### 5.1 Design and Construction

A diagram of the NIST MSMC is shown in figure 6. The housing of the MSMC is composed of two boxes made from 3.18 cm thick 6061 aluminium plates; the boxes are joined one on top of the other but are hermetically isolated from one another. The upper box, which we will refer to as the upper chamber (UC) contains a high precision (0.01 mg resolution) 10 kg mass comparator that has been specially modified for compatibility with a high vacuum environment. The mass comparator can be operated automatically and contains a 4-position turntable for holding mass artefacts, which it may compare independently of any mass in the lower chamber.



**Figure 6. NIST Magnetic Suspension Mass Comparator. Some details of the structural support have been omitted for clarity.**

The magnetic suspension system designed for the MSMC relies on the mutual attraction of upper and lower magnet assemblies, which are shown in detail in figure 7.



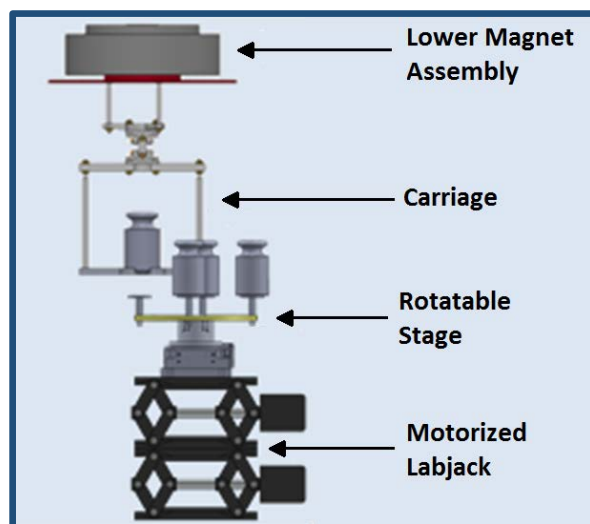
**Figure 7. Detailed diagram of the upper and lower magnet assemblies.**

The upper magnet assembly is located in the upper chamber and is attached to the weighing pan through the dial stack of masses by connecting rods that in turn attach to a suitable adapting plate. In order to accommodate this adapting plate, it was necessary to remove the 5 kg dial weight that is standard equipment for this comparator. There are two other dial masses, 3 kg and 1 kg that remain. The adapting plate in the dial mass stack consists of a circular aluminium plate with two tabs that are spaced 180° apart. Each tab has a hole to accommodate a conical nut that connects to an aluminium down-rod. The other end of the two down-rods connect to the horizontal ends of a mechanically balanced bridge structure that is used to suspend the upper magnet assembly. The upper magnet assembly consists of a Sm-Co cylindrical permanent magnet combined with a solenoid and a soft iron core. The assembly is surrounded by magnetic shielding that shapes and concentrates the magnetic flux lines and prevents loss of magnetic energy due to stray fields. The cylindrical magnet has a flat face and acts as the pole for this assembly. The upper and lower magnet assemblies are separated by an aluminium flange that serves as the boundary between the evacuated upper chamber and the lower chamber which is held at nominally atmospheric pressure. The construction of the lower magnet assembly is similar to that of the upper magnet assembly, using a solid cylindrical Sm-Co magnet as the pole in conjunction with a soft iron core and mu-metal shielding. There is no solenoid on the lower magnet assembly. As shown in figure 6, the lower magnet assembly is attached to a carriage containing a pan that holds a mass artefact. The cylindrical permanent magnets in both upper and lower magnet assemblies are magnetized in the same direction to provide an attractive force between them.

Mass standards are installed in the upper and lower chambers using a specially designed Mass Transport Vehicle (MTV) that incorporates a load-lock (which mates to a similar structure on the upper chamber) and can be evacuated in order to transport and load masses under vacuum conditions. The MTV is a mobile stainless-steel enclosure that is equipped with suitable linear motion actuators and custom-designed holding jigs to capture a mass artefact from the NIST-4 watt balance, transport it to the MSMC, and insert it into the MSCM under

vacuum. All-metal flange seals are used to ensure leak tightness, and a gate valve with pumping port is integrated to allow evacuation or venting prior to admitting an artefact.

In the MSMC upper chamber, up to four 1 kg mass artefacts may be loaded onto the turntable. In the lower chamber, a mass exchange device for loading masses onto and off of the mass carriage has been designed and constructed. This system can also accommodate four mass artefacts that sit on individual pedestals, and installs or removes a mass from the carriage prior to magnetic suspension. The lower chamber masses sit below the carriage, and a motorized lab jack is used to lift the mass of interest to the proper height. A rotating stage then allows the mass to be rotated into position for placement on the carriage. The motorized lab jack is then lowered to place the mass onto the carriage. This process is reversed for removing a mass. A diagram of the lower chamber mass exchange device is shown in figure 8.

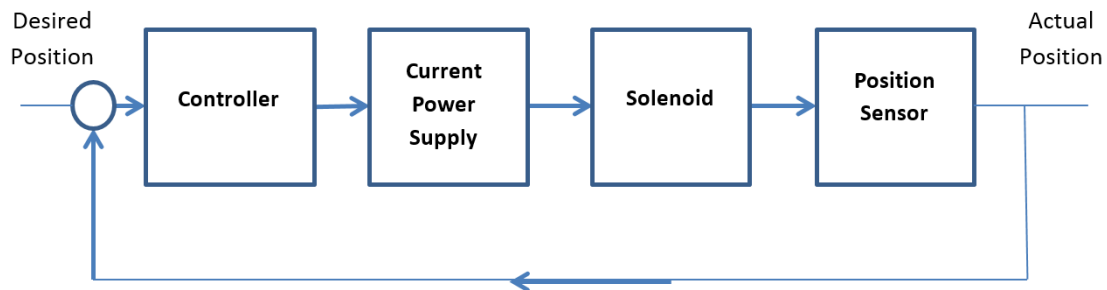


**Figure 8. Lower Chamber Mass Exchange System.**

## 5.2 Magnetic Suspension

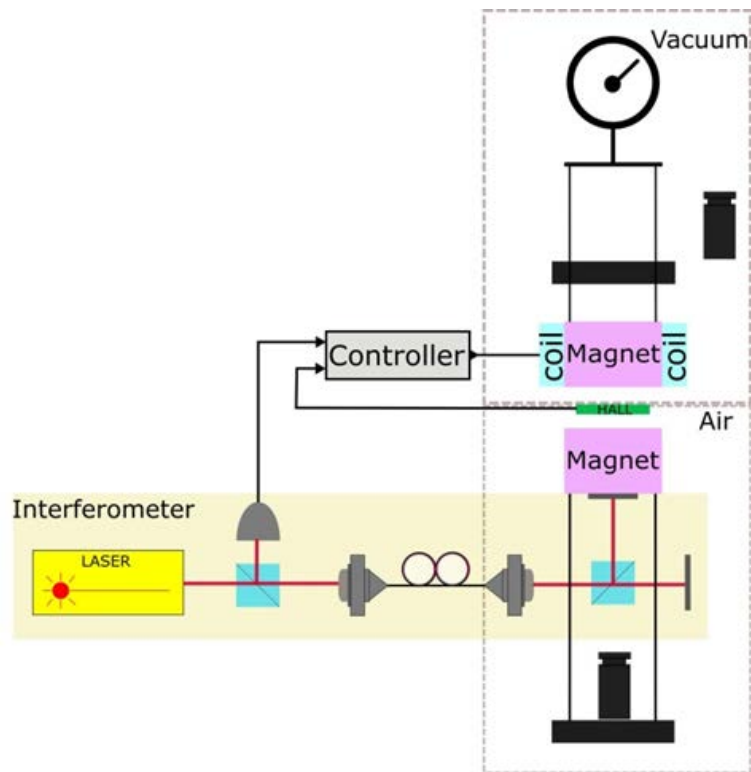
Electromagnetic suspension techniques are covered extensively in the literature [24] and will not be reviewed here. For the MSMC, magnetic suspension occurs when the magnetic force pulling up on the lower magnet assembly plus the carriage is equal to the gravimetric force pulling down on the lower magnet assembly plus the carriage. From this point on, “lower magnet assembly” includes the carriage that is attached to it; there may or may not be a mass on the carriage, but this will be clarified when necessary. The magnetic force on the lower magnet assembly depends nonlinearly on the separation distance of the pole pieces (cylindrical magnets, as shown in figure 5), and it is necessary to position the lower magnet assembly so that the distance between the poles is approximately correct for achieving levitation. This is done by a set of three automated lifting mechanisms, not shown in figure 6. The Hall sensor measures the field associated with this starting position; the servo loop including the Hall sensor and the solenoid current supply maintains this value of the magnetic field by adjusting the current through the solenoid windings in response to vertical

movements up or down of the lower magnet assembly. Figure 9 is a diagram of the suspension servo loop. The solenoid current is provided by a pulse width modulated (PWM) power supply that runs at 100 kHz.



**Figure 9. Magnetic suspension servo loop. The desired and actual positions are voltages that are proportional to the magnetic field between the upper and lower magnet assembly poles**

The entire feedback system resides on a field programmable gate array (FPGA) in order to avoid jitter in the feedback loop caused by computer latency, which would produce instabilities in the magnetic suspension. Furthermore, it was determined through magnetic field simulations that position resolution of the magnetic suspension using the Hall sensor would limit the stability of the suspension to something greater than the resolution of the mass balance. In order to reduce this lower limit, a heterodyne laser interferometer has been incorporated as an additional position sensor that is capable of detecting sub-micron variation in the vertical suspension position. Both the Hall probe signal and the interferometer signal are fed into the servo controller, the output of which drives the solenoid coil in order to stabilize the suspended assembly and improve the precision of the mass measurement. A diagram of the magnetic suspension system including the laser interferometer is shown in figure 10. More details of the suspension system and control can be found in Stambaugh et al [25, 26].



**Figure 10. Schematic of the magnetic suspension system for vacuum-to-air mass comparison.**

### 5.3 Comparison Measurements

The MSMC will be completely automated using software written in LabVIEW. Any weighing sequence can be programmed, for instance A-B-B-A, where A is a mass in vacuum in the upper chamber, and B is a mass in air in the lower chamber. The mass of the carriage in the lower chamber cancels out of the difference of the measurements as the carriage is always suspended regardless of whether a measurement is being made in the upper or lower chamber. If a mass is being measured in the upper chamber, then the mass is first loaded onto the weighing pan in the upper chamber. The carriage (with no mass on it) is then suspended. If it is desired to then make a measurement in the lower chamber, the mass on the upper chamber weighing pan is removed. Then the carriage is taken out of suspension and loaded with the mass of interest. The loaded carriage is then re-suspended and the measurement can begin. For comparison of two masses of the same nominal value, this weighing sequence assures that the same nominal load is on the comparator at all times. It should be noted that the comparison described is different to mass comparison carried out entirely in air (or entirely in vacuum) since air buoyancy will only affect one side of the comparison. This makes knowledge of the determination of air density at the time of comparison a critical measurement since it directly affects the uncertainty in the mass of the standard used in air.

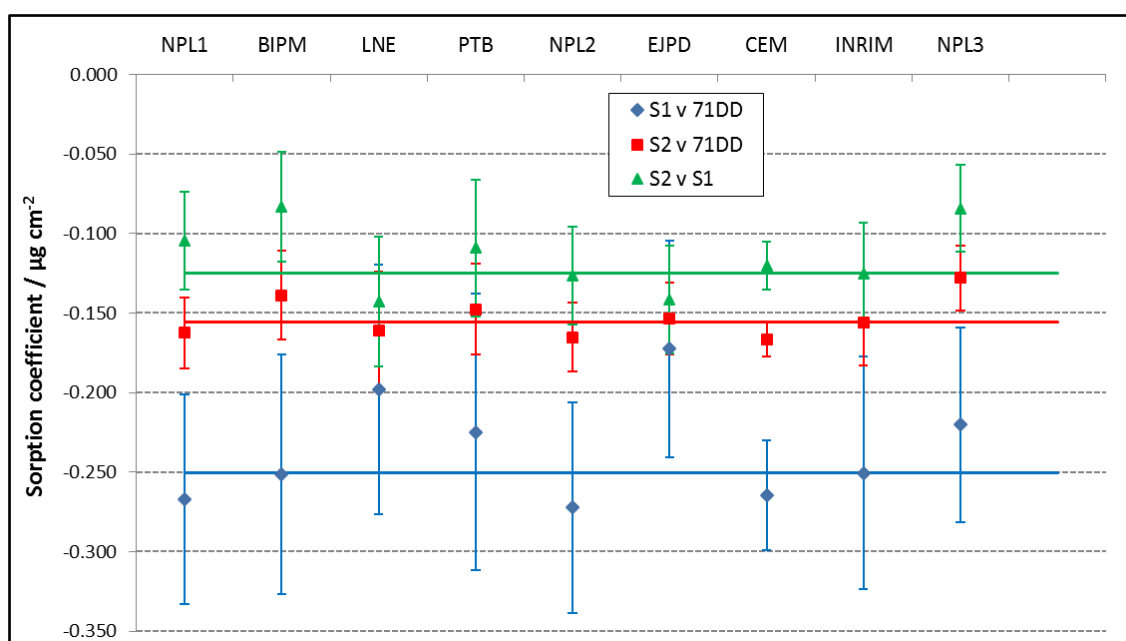
## 5.4 Current Status

As of this writing the NIST MSMC has only been tested in air for stability of the mass reading of a suspended artefact. In experiments using a comparator with 0.1 mg resolution, the stability of the mass reading was limited by the noise of the comparator indication. Furthermore, there is evidence from interferometry measurements that the stability is well below 0.1 mg. A 0.01 mg resolution comparator is currently being installed, which will enable an accurate measurement of the stability of the suspension. Using both the Hall sensor and interferometer signals as feedback to the suspension control loop, we expect the ultimate stability to be 0.02 mg or better. Vacuum-to-air comparisons are expected to commence before summer of 2016.

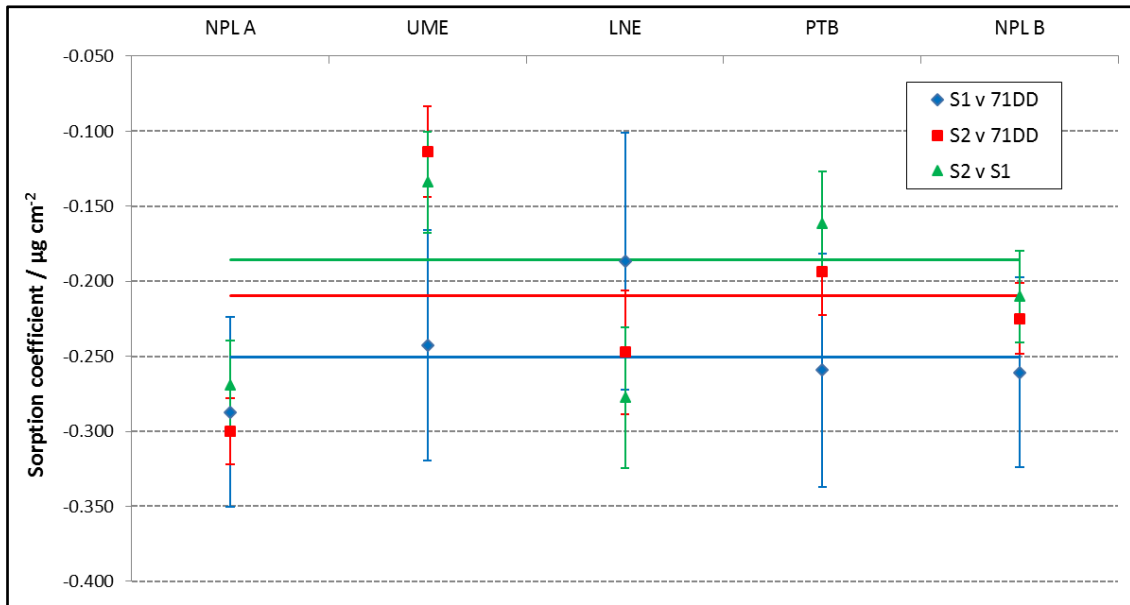
## 6. Comparison data for sorption effects

### 6.1 CCM WGM TG1 comparison of stainless steel sorption artefacts

As part of the work of Task Group 1 (TG1 Mass metrology under vacuum) of the Consultative Committee for Mass and Related Quantities, Working Group on Mass Standards, a comparison of the sorption measured by participating laboratories was undertaken using a set of sorption artefacts as the transfer standards. NPL piloted the comparison and circulated three stainless steel sorption artefacts consisting an integral (cylindrical) kilogram standard (71DD) and two and four piece weights stacks (S1 and S2 respectively) [13]. The transfer standards were circulated in two consecutive petals, the results of the comparison are shown in figures 11 and 12.



**Figure 11. Sorption values determined for the transfer standards in Petal 1, error bars represent standard uncertainties ( $k = 1$ ).**

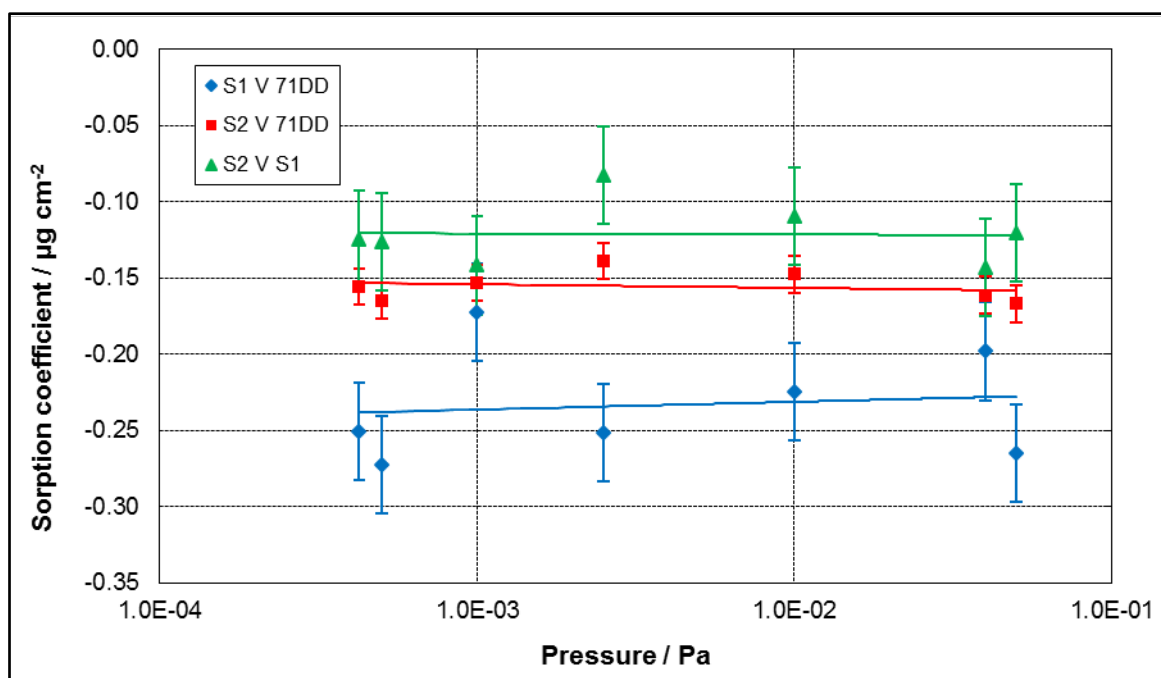


**Figure 12. Sorption values determined for the transfer standards in Petal 2, error bars represent standard uncertainties ( $k = 1$ ).**

The results show good agreement between the participants with regard to the sorption values calculated using pairs of artefacts in different combinations. However, there is a significant and consistent systematic difference in the sorption determined with the different artefact combinations, this is discussed further in section 6.1.1.

### 6.1.1 Sorption pressure dependence

The vacuum pressure level during the sorption measurements was recorded by each participant and the correlation of sorption with pressure level is shown in figure 13.



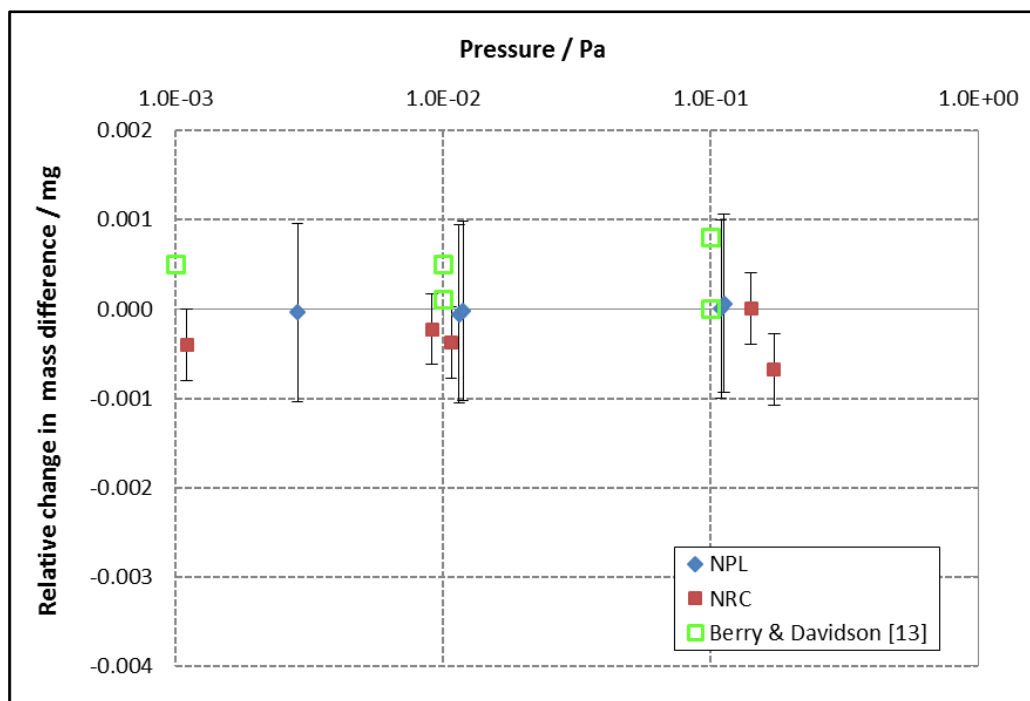
**Figure 13. Participants' sorption values versus vacuum pressure for S1 V 71DD (blue diamonds), S2 V 71DD (red squares) and S2 V S1 (green triangles).**

According to the results shown in figure 13 there does not appear to be any significant correlation between the level of vacuum in the participants' apparatus and the sorption value measured over the range of vacuum pressures used (0.000 43 Pa to 0.05 Pa). This confirmed the earlier results of Berry and Davidson which showed no significant sorption on stainless steel surfaces below a pressure of 0.1 Pa.

The (sorption) results of all the participants can be seen to agree within the estimated uncertainties. This shows that, for a given set of sorption artefacts, the sorption is independent of the vacuum equipment and mass comparator used to make the measurements. However, the sorption values calculated from the comparison of the three artefacts varies depending on which pairs of artefacts are used to calculate the value. The likely source of this discrepancy is either an error in the measured volumes of the weights or an inhomogeneity in the surfaces of the artefacts. Because the difference between the mass values of the artefacts needs to be determined in air and in vacuum the measured (differences in the) volumes of the artefacts is a major contribution to the uncertainty in the sorption value calculated. However, the volumes of the artefacts were re-determined between Petal 1 and Petal 2 of the comparisons and the values used for the comparisons were found to be correct. Thus, it was assumed that the differences in the pair-wise sorption values determined were due to differences in the (behaviour of the) surfaces of the artefacts. Evaluation of such discrepancies is discussed in Section 9.

## 6.2 Comparison of sorption pressure dependence

Based on the results of the comparison and the findings of Berry and Davidson a pressure range of 0.1 Pa to 0.001 Pa is recommended for weighing in vacuum, and for the operation of watt balance experiments. A comparison was undertaken to confirm the uniformity of surface sorption over this pressure range for stainless steel, platinum-iridium and silicon kilogram artefacts. The results of bilateral comparisons between NPL and NRC and between INRIM and CMI are shown in figure 14 to figure 16. For each participant the gravimetric sorption effects at lower pressure are reported relative to their sorption at value 0.1 Pa. Previous results for the same artefacts, reported by Berry and Davidson [14] are also given.



**Figure 14. Change in mass difference between the NPL Pt-Ir artefacts measured at NPL and NRC relative to the initial mass difference at 0.1 Pa. The artefacts were cleaned before both sets of measurements and error bars represent the standard uncertainty ( $k = 1$ ). Open markers represent previously published values.**

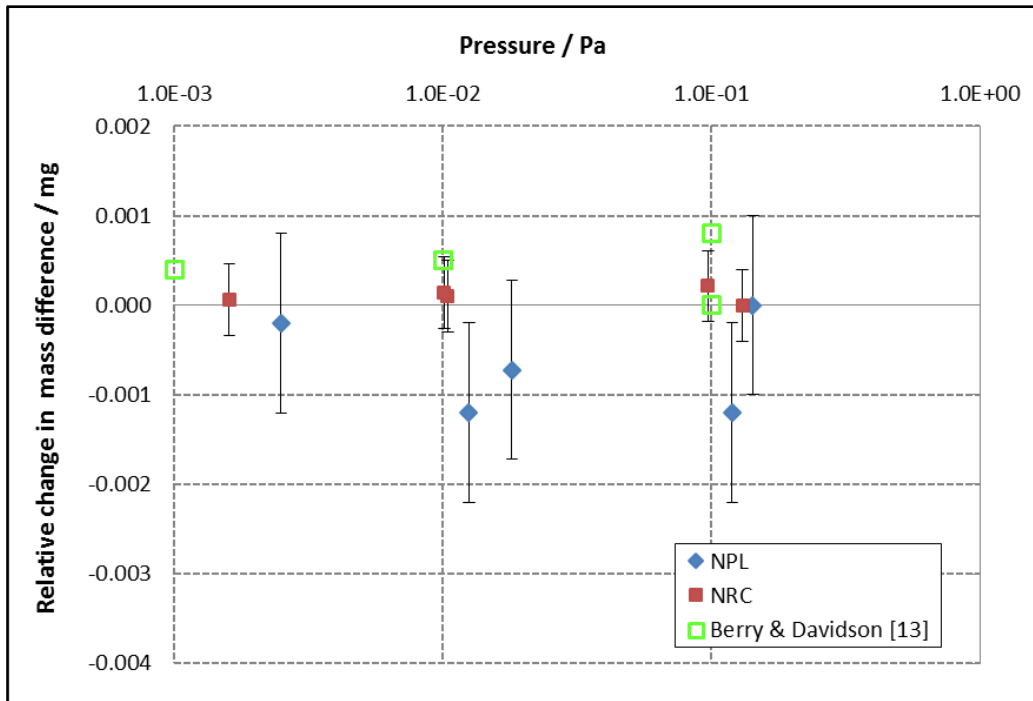


Figure 15. Change in mass difference between the NPL silicon artefacts measured at NPL and NRC relative to the initial mass difference at 0.1 Pa. The artefacts were cleaned before the NPL measurements. Error bars represent the standard uncertainty ( $k = 1$ ). Open markers represent previously published values.

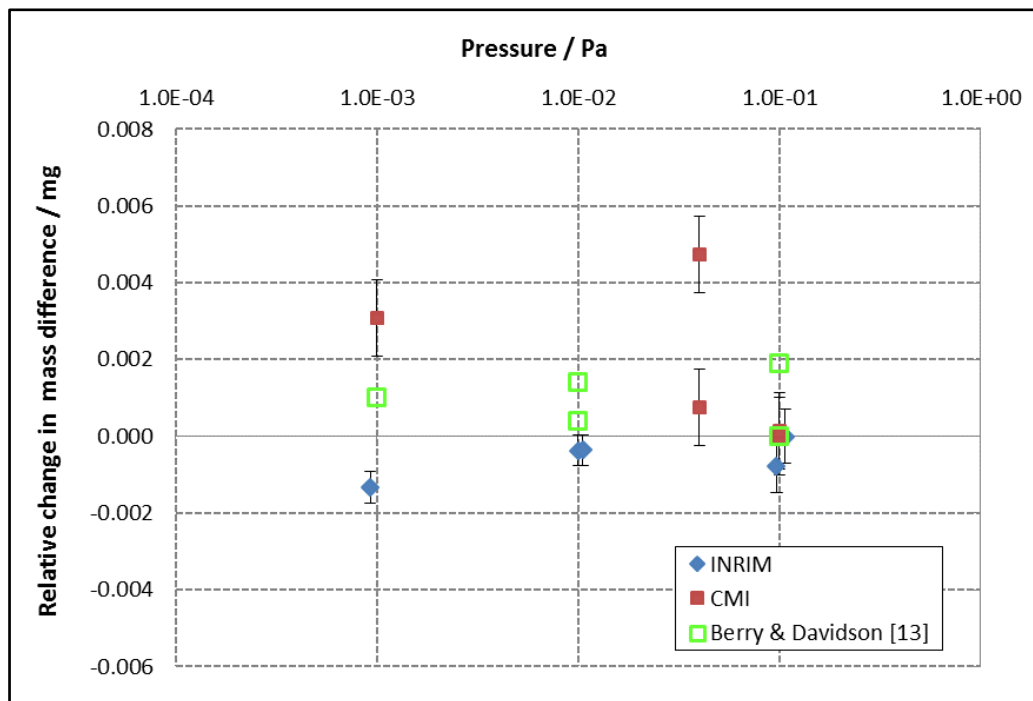
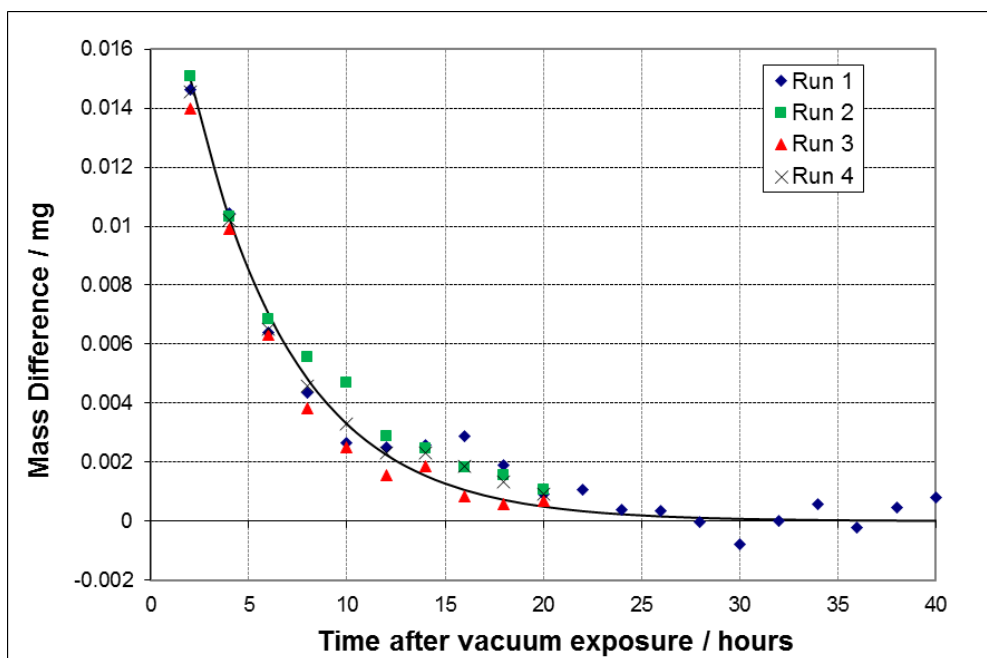


Figure 16. Change in mass difference between the INRIM stainless steel artefacts measured at INRIM and the stainless steel artefacts measured at CMI relative to the initial mass difference at 0.1 Pa. The artefacts were cleaned before the measurements and error bars represent the standard uncertainty ( $k = 1$ ). Open markers represent previously published values.

With the exception of the CMI data, which show a rather large variation, the results of the comparisons confirm the pressure range of 0.1 Pa to 0.001 Pa can be recommended as suitable for the measurement in vacuum of masses used in the dissemination of the mass scale, the gravimetric effect on mass standards used in this range being less than 2 µg for all measurements.

## 7. Desorption dynamics

Figure 17 shows the results of gravimetric desorption measurements made at NPL on a stainless steel kilogram moved from air to vacuum (pressure 0.001 Pa). The transfer was carried out 4 times to determine the repeatability of the process, the mass difference being relative to a kilogram standard which had been stored in vacuum for 1 month.



**Figure 17: Gravimetric determination of desorption from a stainless steel kilogram moved from air to vacuum (pressure  $10^{-3}$  Pa). The nominal standard uncertainty of each measurement is  $\pm 1.0$  µg.**

The results show that the desorption of mass from the surface of the weight follows a very similar pattern for all 4 measurements. The initial weighing was made about 2 hours after the introduction of the weight into vacuum so the total (air to vacuum) mass change is not apparent but was of the order of 20 micrograms. The desorption follows an exponential decrease of the form;

$$\Delta m = \Delta m_{\text{fin}} e^{-a.t} \dots\dots\dots(3)$$

Where  $\Delta m$  is the mass loss at time  $t$  after initial vacuum exposure (in hours),  $\Delta m_{\text{fin}}$  is the total mass loss for air-vacuum transfer, and  $a$  is a constant relating to the desorption of the surface of the mass standard (in the case of the weight used for this test,  $a = 0.19$ ). The results show that a period of 24 hours is sufficient to obtain a stable value for a mass standard transferred to vacuum from ambient air.

At NRC additional surface analysis measurements were made on surfaces transferred between air and vacuum. A typical desorption process at a vacuum pressure of  $10^{-5}$  Pa is shown in figure 18a. Data was acquired by X-ray photoelectron spectroscopy (XPS) on a Kratos Axis Ultra spectrometer and shows the desorption of carbonaceous species as a function of time after introduction to vacuum. The nominal values and error bars represent the average and standard deviation respectively of 3 distinct measurement areas on a single sample of PtIr foil. While water is expected to represent the major portion of total sorption change between air and vacuum, the majority of the mass in the form of water is likely to be removed at higher pressures during pump-down (as can be seen in figure 5). For PtIr, water remaining at  $1 \times 10^{-5}$  Pa is strongly bound to the metal (or metal oxide) surface or the carbonaceous overlayer. Analysis of the C1s peak envelope shows that it is largely hydrocarbon bonds ( $\sim 50\%$  of the overlayer), but it also contains carbonyls, alcohols, and other oxygen containing functional groups. Figure 18a shows the surface as a function of time for the total C1s envelope with the desorption following kinetics in which the rate is proportional to the surface coverage. Fitted according to the exponential decrease with time, it takes a similar form as shown in figure 17 though the rate is slower with the desorption of hydrocarbons on the PtIr surface studied showing an exponential dependence of  $0.049(t-t_0)$ . After approximately 24-40 hours of vacuum exposure it can be seen that the desorption rate reduces to near zero indicating the approximate stabilisation time at this level of vacuum. It should be noted that the desorption rate reaches zero at non-zero layer thickness ( $\sim 1.3$  nm), so that surfaces with cleaner initial state on introduction to vacuum have the potential to increase their initial overlayer thickness and therefore to gain mass. This steady state surface coverage is not only governed by the properties of the metal and contaminant overlayer but also of the residual gas content in the vacuum chamber since sorption is a combination of adsorption and desorption processes.

The desorption curve by XPS in figure 18a agrees in form with gravimetric observations for PtIr (see figure 18b), however unlike the XPS results and the vacuum cycling of figure 17, figure 18b is a differential measurement between two PtIr artefacts, which are changing dynamically together during pump-down. The result is only a small residual change less than  $1 \mu\text{g}$ . XPS can be considered an absolute measurement in which the magnitude of the detected signal is approximately proportional to the quantity of surface atoms. Gravimetric measurements on the other hand are performed in a comparator and are relative measurements so only the difference in sorption kinetics between compared masses will be observed. When weighing masses with similar surface characteristics and surface areas, in theory no kinetic trends would be observed from a (gravimetric) comparison in a vacuum balance, unless the masses compared had been stored in vacuum for different periods. Since, like XPS, watt balances are absolute measurement devices any kinetic change occurring in vacuum would be observable in the form of a drift in the measured mass value.

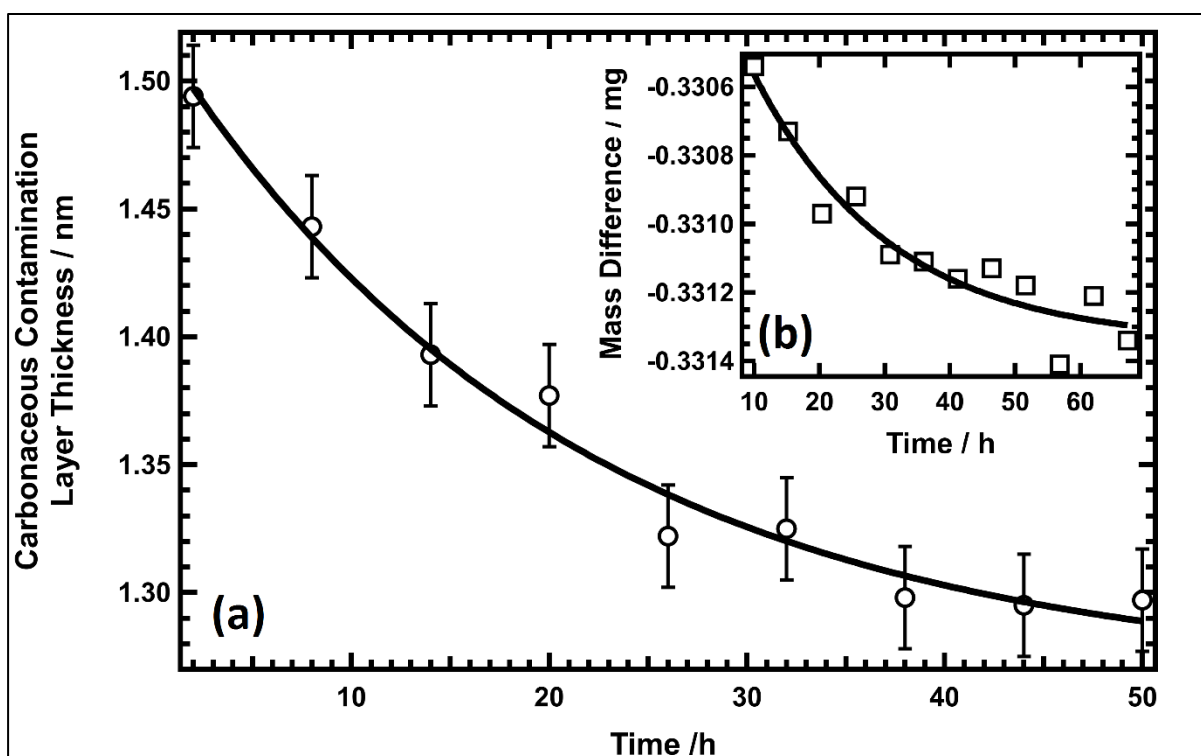


Figure 18: Carbonaceous overlayer thickness on PtIr surface measured by XPS at successive times after vacuum exposure at  $1 \times 10^{-5}$  Pa. The error bars are derived from the standard deviation of 3 measured locations. The inset (b) shows the similar kinetics for gravimetric measurements between two Pt-Ir artefacts (0.0002 mg nominal uncertainty).

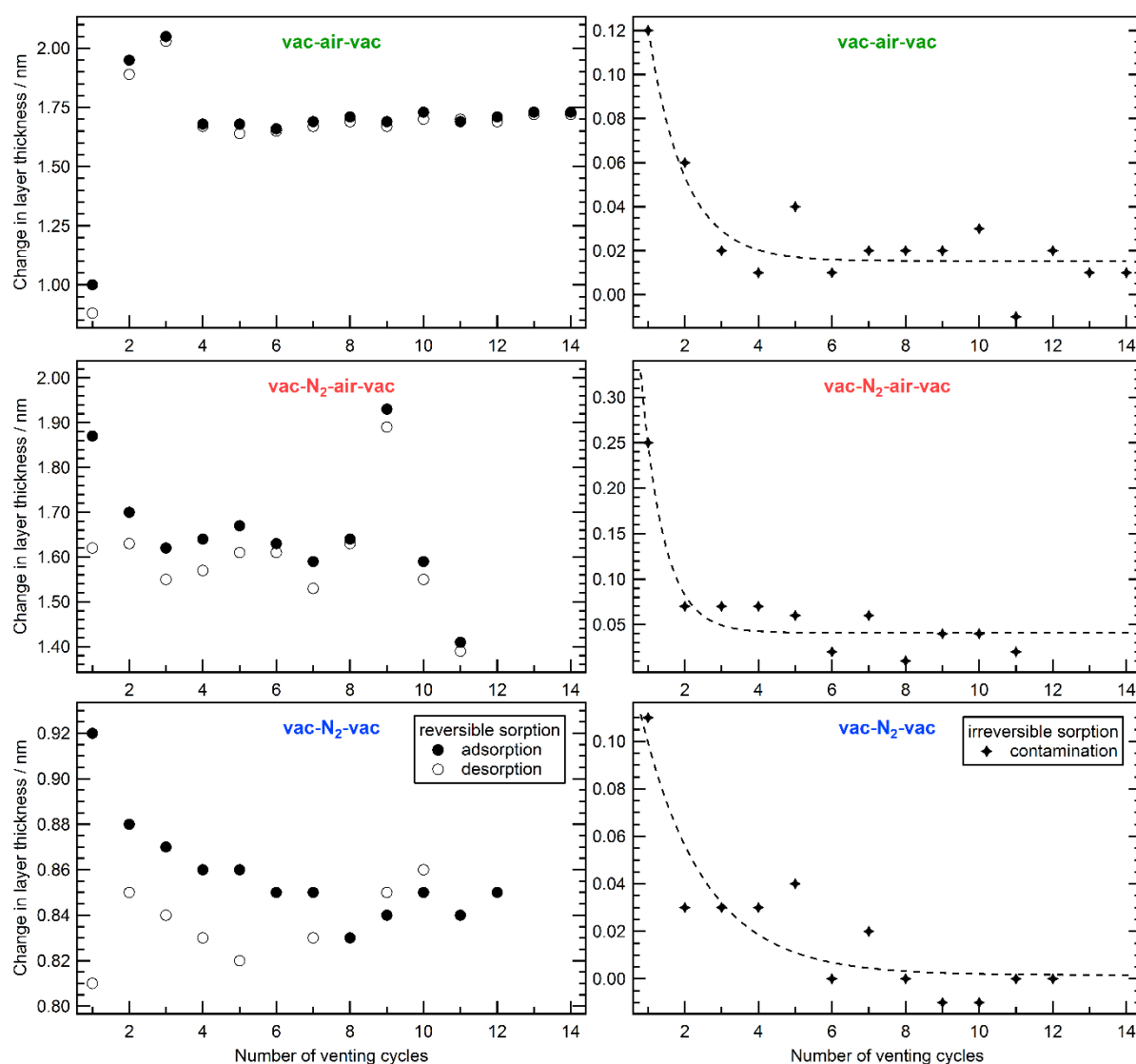
## 8. Evaluation of venting cycles.

Previous work by Davidson [27] noted the accelerated accretion of hydrocarbon contamination on platinum-iridium mass standards transferred between air and vacuum when compared with similar standards stored permanently in either of the two media. Davidson's work used both XPS and gravimetric measurements to monitor the stability of mass standards. Recently experiments have been undertaken at METAS and at NPL to evaluate the effect of different venting cycles on the stability of the transfer standards. The venting cycles used were;

1. Vacuum-Air-Vacuum
2. Vacuum-Nitrogen-Vacuum
3. Vacuum-Nitrogen-Air-Vacuum

The use of nitrogen to passivate the (baked) surfaces of vacuum chambers before air exposure is well known [28]. The use of nitrogen as an intermediate stage in vacuum air transfer has been evaluated to quantify the potential beneficial effect on maintaining the cleanliness of the surfaces of mass transfer standards. Work undertaken by METAS studied the recontamination of surfaces following plasma cleaning [29] for the three venting cycles using

a quartz crystal microbalance (QCM) with gold plated crystals. The results are shown in figure 19.

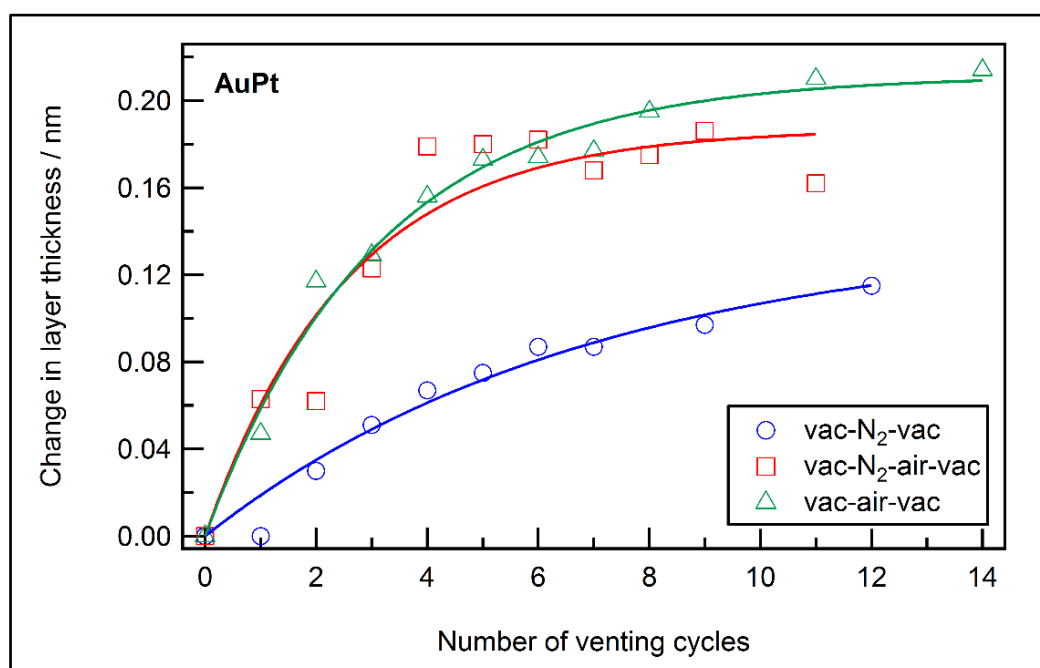


**Figure 19: Sorption and desorption measurements using a gold-coated quartz crystal microbalance: The left hand graphs show reversible sorption (adsorption: dots and desorption: circle) and the right hand graphs show irreversible sorption (contamination: star) for three different venting cycles.**

The results show that the cleaned surfaces accrete significant contamination (irreversible sorption) on the first venting cycle with the vac-N<sub>2</sub>-air-vac cycled sample gaining the most. Subsequent cycles show much smaller levels of irreversible sorption. The total reversible sorption is approximately the same for the two air-vacuum cycles with the intermediate nitrogen step not affecting the amount of water sorbed onto the surface. However, the repeatability of the vac-N<sub>2</sub>-air-vac cycled sample is much poorer than that of the sample

cycled only between air and vacuum. The vacuum-nitrogen cycled sample shows much lower surface (water) sorption levels, which is to be expected since nominally dry nitrogen was used.

METAS also conducted similar tests using samples of various metals under investigation as candidate materials for new mass standards (gold, gold-platinum alloy, platinum-iridium, pure iridium and copper samples plated with rhodium and with gold). The (non-reversible) surface contamination of the samples was monitored using XPS. The results for the tests showed similar behaviour to that of the QCM crystals. The results for the gold-platinum alloy sample are typical of the results for all the test samples and are shown in figure 20.

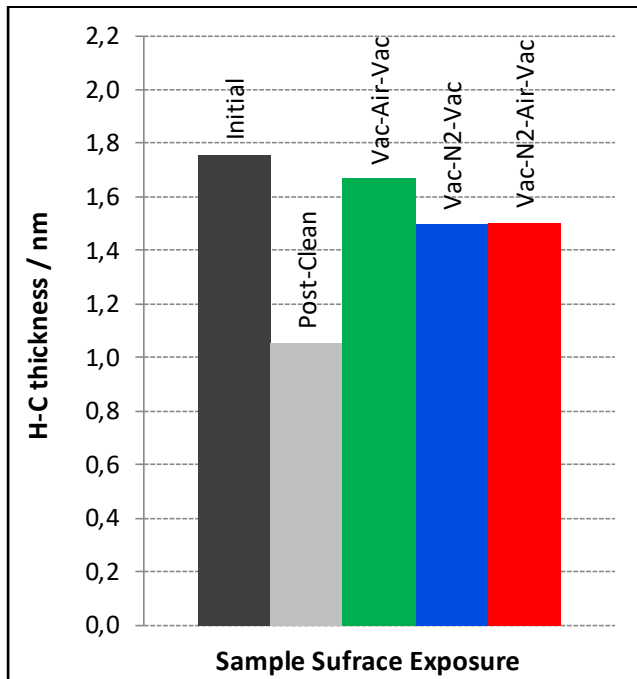


**Figure 20: The overlayer thickness of hydrocarbon on gold-platinum alloy is calculated from XPS measurements for cyclic venting between vacuum and air and/or nitrogen. The data are fitted by the self-limited growth model [30, 31]**

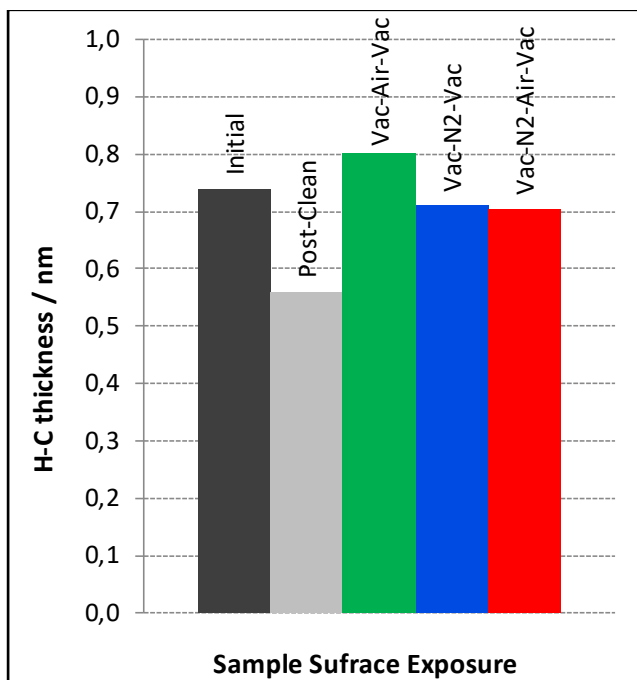
The results show that the increase in the contamination layer is similar for both the air-vacuum cycles and that the inclusion of the intermediate nitrogen stage does not significantly inhibit overlayer growth. After approximately 4 cycles the increase in layer thickness approaches a self-limited value. The value for the gold-platinum sample is approximately 0.2 nm and was between 0.1 nm and 0.3 nm for all the surfaces tested. The vacuum nitrogen sample remains cleaner but it may be supposed that with further cycling the overlayer will also approach the self-limiting value.

NPL also conducted an evaluation of the effect of different venting cycles on the growth of the contamination layer using XPS. Samples of platinum-iridium and silicon were cleaned using the UV-Ozone cleaning process [32] and then cycled five times between vacuum and air or nitrogen using the three venting cycles described previously. The results of the tests are given in figures 21 and 22. Three samples of each material were used, the initial and post-

clean results represent the average value for the three samples. The variation in these values was less than 0.1 nm in all cases.



**Figure 21. NPL results for platinum-iridium samples exposed to the three venting cycles identified. Initial and post clean values represent the average for all three samples.**



**Figure 22. NPL results for silicon samples exposed to the three venting cycles identified. Initial and post clean values represent the average for all three samples.**

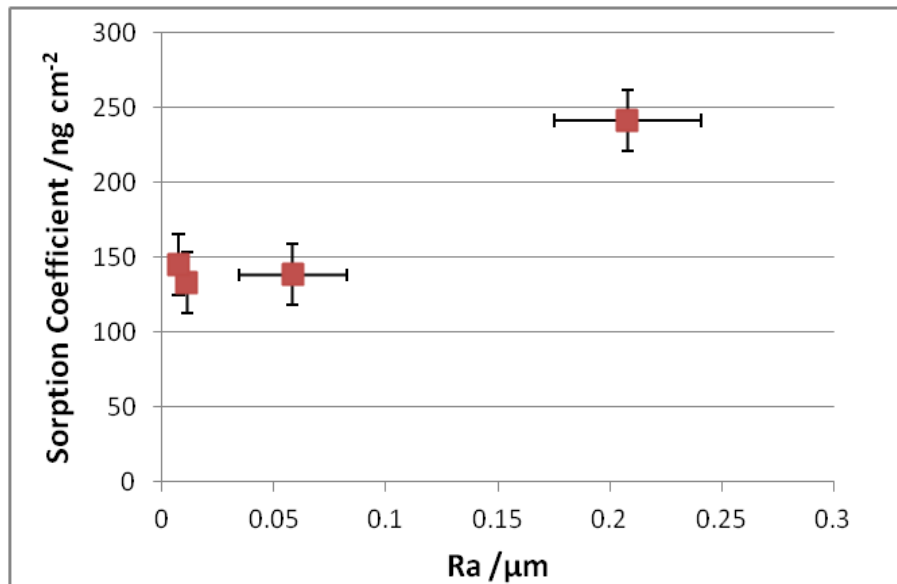
In comparison with the METAS results the NPL measurements show that both the Vac-Nitrogen-Vac and Vac-Nitrogen-Air-Vac cycles give a slight improvement in restricting the recontamination of the samples but the difference to the Vac-Air-Vac results is not very significant (about 0.1 nm for both materials). It is likely that as the number of cycles is increased the re-contamination of all samples would approach the self-limited value as indicated by the METAS results.

## **9. Sorption correlation with surface roughness**

Sorption values in table 1 show that there is as much variation in sorption between artefacts manufactured from the same material as there is between artefacts made from different materials. This suggests that sorption values are not influenced as much by the bulk material the artefacts are made from, but by other factors. Equations 1 and 2 estimate the sorption coefficient per unit surface area with the assumption that the average roughness of the integral artefact is equal to that of the stack(s). This assumption is an over simplification since perfect homogeneity in surface roughness of the integral and stack of the sorption artefact pair is not easily obtained during the polishing process, and even more difficult to maintain during use.

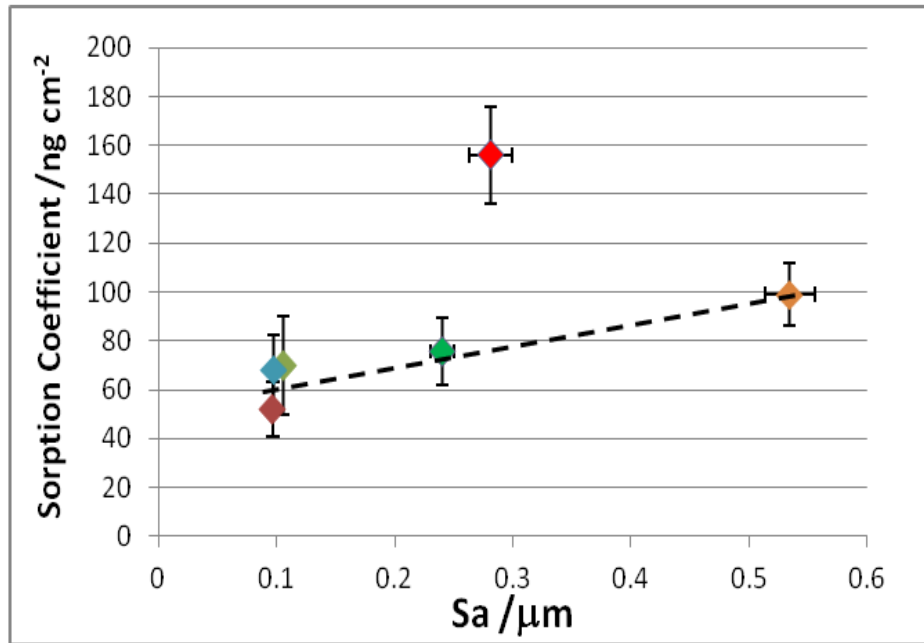
Studies have been undertaken at INRIM and NPL to examine the correlation of surface roughness with sorption between air (50 % relative humidity) and vacuum ( $p < 0.1$  Pa) on artefacts of varied surface quality and material. In the study by NPL masses were not cleaned before testing, in the study by INRIM all masses were cleaned. INRIM examined one platinum-iridium and four stainless steel artefacts, with sorption coefficients ranging from  $0.05 \mu\text{g cm}^{-2}$  to  $0.16 \mu\text{g cm}^{-2}$ . NPL measured four stainless steel artefacts, with sorption coefficients ranging from  $0.13 \mu\text{g cm}^{-2}$  to  $0.24 \mu\text{g cm}^{-2}$ .

The roughness parameters were measured using two different variants of profilometry; 3D profilometry at INRIM and 2D profilometry at NPL. To obtain a value representative of the whole artefact, roughness was measured at different points on the surface. The 3D measurements were performed at 25 sampling areas of  $1.00 \text{ mm} \times 1.27 \text{ mm}$  each, whereas the 2D surface roughness parameters were measured at 3 profiles approximately 8 mm long. The standard deviation of the measurements are given as the measurement uncertainty, though probe shape, dynamic response and sampling will all contribute to the measured roughness value. Therefore direct comparison of absolute roughness values between the 2D and 3D profilometry should be treated cautiously. In the case of the 2D profilometry the two parameters  $R_a$  and  $R_z$  [33] were analyzed; in the case of the 3D profilometry, among the many 3D parameters available, the most suitable to describe the influence of the surface quality on the sorption were  $S_a$ ,  $S_z$ ,  $S_q$ ,  $S_v$ ,  $S_{dr}$  and  $S_{10z}$  [34]. However, for both approaches, no particular roughness parameter emphasized a stronger correlation with sorption than another. The results obtained at NPL with the  $R_a$  parameter and at INRIM with the  $S_a$  ( $R_a$  parameter measured over an area) are shown in figures 23 and 24, respectively.



**Figure 23: NPL measurements of uncleaned stainless steel mass artefacts with roughness values determined via 2D profilometry. The horizontal error bars represent the standard deviation of 3 measurement profiles.**

From the NPL results, only the weight with the highest roughness (OIML F1/F2 equivalent) [35] shows significantly different sorption characteristics. The other artefacts with roughness values measured to be compliant with OIML recommendations for Class E1 weights, show no significant variation in sorption. This supports the conclusion by Schwartz [36] that the sorption behavior of OIML Class E1 masses cannot be altered significantly by improving their surface roughness further (OIML E1 specifies  $R_a < 100 \text{ nm}$ ). From Kochsiek [37] it was shown that a tenfold increase in roughness produces a  $100 \text{ ng/cm}^2$  increase in sorption (over the range  $0.1 \text{ } \mu\text{m}$  to  $3 \text{ } \mu\text{m}$   $R_z$ ). For the NPL results, in which  $R_z$  scales approximately as  $10 \times R_a$ , a similar behavior is observed, the sorption difference between artefacts with  $R_z$  roughness in the range  $0.07 \text{ } \mu\text{m}$  to  $2.2 \text{ } \mu\text{m}$  is also approximately  $100 \text{ ng/cm}^2$ .



**Figure 24: INRIM measurements of sorption difference on pre-cleaned artefacts of stainless steel and PtIr, with Sa roughness as determined by 3D profilometry. The horizontal error bars represent the standard deviation of 25 measurement profiles. The dashed line is a linear fit excluding NPL SS.**

The artefacts of the INRIM study (figure 24) span a larger range, from 100 nm to over 500 nm (Sa). With the exception of artefact NPL SS, a linear correlation is apparent and the sorption values are significantly lower than those observed in the NPL study. This difference is probably due in part to the surface cleanliness of the artefacts. In the work of Schwartz [10, 36] uncleaned masses showed about 2.5 times greater sorption than cleaned ones, approximately the difference between the sorption reported in the NPL and INRIM studies. The origin of the difference in behavior of NPL SS artefact is not known but is possibly due to surface scratches or an increased depth in the oxide layer of these artefacts. The NPL SS artefact used in the INRIM study was a disc of stainless steel, part of a stack used in the CCM WGM TG1 comparison [13]. The unexpected value may be due to surface damage or other defects in locations of the disc which were not analyzed during the roughness measurements or are not observable by profilometry. Additionally, sorption is known to be dependent not only on cleanliness but also oxide thickness and identity [37]. The influence of the bulk material and surface preparation, which can change the oxide thickness and chemical makeup, will also play a role. The masses in the INRIM study were each acquired at different times from different manufacturers and therefore these influence parameters could not be controlled.

Based on the results of this study, it is not possible to establish a firm correlation between the sorption coefficient and the surface roughness, however it is clear that particular attention should be paid to the roughness of the sorption artefacts especially when Ra roughness exceeds 100 nm (OIML E1). The contact surfaces of stacked disc weights are subjected to scratches, oxide formation and contamination, all of which will increase the sorption. Without

Careful control during use the sorption coefficient per unit surface area for the stacks and integral could deviate leading to significant errors in the sorption determination using artefacts.

# 10. The robust determination of sorption values from gravimetric data and uncertainty calculations

In order to be able to validate the transfer of the unit of mass between air and vacuum, it is recommended to follow a procedure, where at least one mass standard is permanently stored in vacuum and one mass is permanently stored in air. Although a set of two sorption standards are sufficient in principle, it is recommended to use a set of at least three sorption standards, which are cycled between air and vacuum a number of times. In each cycle the differences in mass of the sorption standards and those kept in air and in vacuum should be measured. Assuming that the masses of the standards permanently stored in air and vacuum are stable, the cyclic repetition allows any contamination of the sorption standards taking place during the cycle to be quantified. The use of sorption standards for the air-vacuum mass transfer is based on the assumption that all sorption standards experience the same desorption/adsorption (including any contamination) per unit surface area during the transfer. In order to test the validity of this assumption, at least three sorption standards should be used.

Assume that a sorption standard  $S$  is cycled repeatedly between air and vacuum. Initially, the standard  $S$  is in air and has mass value  $m_{S,0}$ . When transferred to vacuum for the first time, water and other molecules desorb from the surface, and as a result the standard  $S$  has a new, lower mass value  $m_{S,1}$ . When transferred back to air, some contaminating molecules might be adsorbed to the surface before being protected by an adsorbed layer of water, and as a result the standard  $S$  has a new mass value  $m_{S,2}$ , which could be either smaller or larger than the initial mass value  $m_{S,1}$ . Let  $A_S$  be the surface area calculated from the geometry of the sorption standard  $S$ . Assuming that all parts of the surface of the sorption standard experience the same sorption, the change in mass due to the sorption taking place in transfer no.  $i$  between air and vacuum may then be modelled by

$$m_{S,i+1} = m_{Si} + s_{S,i}A_S \dots \dots \dots (4)$$

where  $s_{S,i}$  is a sorption coefficient (mass of adsorbed/desorbed layer per unit area), which is specific for the particular surface of the sorption standard  $S$  but might vary from one transfer between air and vacuum to the next. When manufacturing a set of sorption standards  $S_1, \dots, S_N, N \geq 2$ , their surfaces might have different sorption efficiencies (e.g. due to differences in surface roughness) leading to different sorption coefficients  $s_{S_1,i}, \dots, s_{S_N,i}$  for the sorption standards  $S_1, \dots, S_N$  in a specific transfer (no.  $i$ ) between air and vacuum. In terms of an average sorption coefficient  $s_i$  for transfer no.  $i$ , this variation may be modelled by

$$\begin{aligned} s_{S_1,i} &= e_{S_1,i}s_i, \\ &\vdots \\ s_{S_N,i} &= e_{S_N,i}s_i, \end{aligned} \dots \dots \dots (5)$$

where the sorption efficiency factors  $e_{S_1}, \dots, e_{S_N}$  fulfils the constraint

$$e_{S_1} + \dots + e_{S_N} = 1 \dots\dots\dots(6)$$

Under the assumptions that all surfaces are almost identical, the best estimates of the sorption efficiencies are

$$e_{S_1} = \dots = e = 1 \dots\dots\dots(7)$$

but with standard uncertainties

$$u(e_{S_1}) = \dots = u(e_{S_N}) = u(p) > 0 \dots\dots\dots(8)$$

In order for the sorption standards to be useful, the standard uncertainty  $u(e)$  should ideally be in the interval

$$0 \leq u(e) \leq 0.1 \dots\dots\dots(9)$$

The resulting model for the change in mass of the sorption standards,

$$m_{S,i+1} = m_{Si} + e s_i A_{S,i} \dots\dots\dots(10)$$

has to be combined with the models for comparing the masses of the sorption standards with the mass kept in air and that kept in vacuum, respectively. The model for the mass comparison in air includes the air buoyancy corrections in terms of the air density and the volumes of the sorption standards and the mass kept in air. In order to be able to separate changes in sorption from changes in air buoyancy, the uncertainties in the volumes of the sorption standards should be as small as possible. By performing several cycles of mass comparisons in air and vacuum, the assumed stability of the masses permanently stored in air or vacuum and the assumed model for the change in mass of the sorption standards during the cycle can be tested using the method of least squares followed by a test of consistency between data and model.

Numerical simulations indicate that the procedure described above allows the unit of mass to be transferred from air to vacuum (or reverse) with an associated standard uncertainty component of less than 0.006 mg. The simulations also indicate that if the assumptions made are not fulfilled in reality, there can be a significant error in the transferred mass value that might not be detected by the consistency test in a single experiment with a limited number of cycles. The reliability of the mass transfer using a particular set of sorption standards should therefore be evaluated by repeating the experiment several times using the same standards permanently kept in air and in vacuum.

## 11. Discussion

A considerable amount of data has been published with regard to the effect of air-vacuum transfer on the value of mass standards. While the phenomenon of the sorption of water on the surface of the transfer standards is ostensibly well understood it is clear that there is

considerable variation in the magnitude of the air-vacuum mass change and in the amount of water adsorbed or desorbed from the surfaces of the transfer standards. Factors such as artefact material and surface roughness and cleanliness have been shown to affect the magnitude of the surface (water) sorption but these alone cannot fully explain the variation in the sorption phenomena seen in the published results. It is likely that additional factors such as the depth and composition of the surface oxide layer also have a significant effect on the amount of water adsorbed. Further work looking at the parameters which affect surface sorption would be beneficial in order to optimise the material and in particular the finishing process used in the manufacture of future transfer standards for the establishment of traceability between mass in vacuum and in air.

Comparisons using a set of surface sorption artefacts as described have shown good repeatability between vacuum balance equipment at different NMIs and have allowed an optimum operating level of vacuum to be recommended. Uncertainty contribution of less than 6 micrograms for the transfer of mass between air and vacuum is achievable using a set of sorption artefacts but care needs to be taken that potential variations in surface sorption between the components of the set are fully assessed and accounted for as described in section 10 of this paper. Further practical work in this area will ensure that a robust uncertainty estimate can be achieved.

An alternative to the use of sorption artefacts is the direct linking of mass in vacuum and air via the Magnetic Suspension Mass Comparator described in section 5. This eliminates the sorption variability issues described but is currently limited in accuracy mainly by the repeatability of the magnetic coupling but also by the need to support a larger tare load than is normal for a kilogram mass comparator. Further development in the equipment will reduce the uncertainty and will provide a useful alternative vacuum air transfer process.

## **12. Conclusions and recommendations**

The transfer of mass between vacuum and air will be necessary in order to disseminate the unit of mass following its redefinition in 2018. To minimise the additional uncertainty contribution to the dissemination, the transfer process between vacuum and air needs to be well characterised and repeatable. The following recommendations are made in order to achieve this;

1. An operating pressure range of 0.1 Pa to 0.001 Pa is recommended for vacuum comparators and watt balance experiments. No significant change in surface sorption has been seen over this pressure range so the compatibility of different experiments can be optimised.
2. Real time sorption measurements should be made for transfer standards rather than using extant or published data. This will minimise the uncertainty contribution from variations in surface sorption effects.
3. The direct transfer of artefacts between vacuum and air is recommended. The use of an intermediate nitrogen stage has been shown to have little benefit to the repeatability of the transfer process.

4. When using sorption artefacts maximise surface area ratios to optimise the accuracy of the sorption calculation.
5. The use of more than two artefacts in a set will allow a more robust assessment of the uncertainty in the surface sorption of the transfer standard. The assessment of the repeatability of the transfer process is also recommended as described in section 10 of this paper.
6. A surface roughness ( $R_a$ ) of better than  $0.1\ \mu\text{m}$  should be realised for sorption artefacts to minimise the magnitude and optimise the repeatability of sorption effects. There is little variation of surface sorption with surface finish provided  $R_a$  is less than  $0.1\ \mu\text{m}$ .
7. The surface quality of the sorption artefacts, and in particular the discs, should be closely monitored due the likelihood of scratching due to the stacking of the discs.
8. The material and finish techniques used to produce sorption artefacts for air vacuum transfer should be considered particularly with respect to the potential thickness of the surface oxide layer produced and the surface hardness (scratch resistance).
9. Determination of the density of sorption artefacts is usually the most significant uncertainty component when determining vacuum-air sorption effects with artefacts and ideally uncertainties of  $0.001\ \text{cm}^3$  or better in the volume differences between artefacts should be achieved.
10. Methods such as the Magnetic Suspension Mass Comparator can provide alternative traceability route to directly achieve the link between masses in vacuum and in air. Here the determination of the density of the air at the time of the (vacuum-air) comparison is a critical measurement in assigning a mass value to the weight in air.

## References

- [1] Steiner R., History and progress on accurate measurements of the Planck constant, *Rep. Prog. Phys.*, 2013, **76**, 1-46
- [2] Becker, P, History and progress in the accurate determination of the Avogadro constant, *Rep. Prog. Phys.*, 2001, **64**, 1945–2008
- [3] Sanchez C.A., Wood B. M., Green R. G., Liard J. O., Inglis D., A determination of Planck's constant using the NRC watt balance, *Metrologia*, 2014, **51**, S5–S14
- [4] Azuma Y., et al, Improved measurement results for the Avogadro constant using a  $^{28}\text{Si}$ -enriched crystal, *Metrologia*, 2015, **52**, 360–375
- [5] SI Brochure: The International System of Units (SI) [8th edition, 2006; updated in 2014] [http://www.bipm.org/utils/common/pdf/si\\_brochure\\_8.pdf](http://www.bipm.org/utils/common/pdf/si_brochure_8.pdf)
- [6] Barrett H. M., Birnie A. W., Cohen M., The Adsorption of Water Vapor on Silica Surfaces, by Direct Weighing, *J. Am. Chem. Soc.*, 1940, **62** (10), 2839–2844
- [7] Masui R. et al, Densimetry in compressed fluids by combining hydrostatic weighing and magnetic levitation *Rev. Sci. Instrum.*, 1984, **55**, 1132-1142
- [8] Wagner W and Kleinrahm R, Densimeters for very accurate density measurements of fluids over large ranges of temperature, pressure, and density, *Metrologia*, 2004, **41**, S24-S39

- [9] Kerl K., A novel weighing method using magnetic coupling of balance and sample by a permanent magnet, *J. Phys. E: Sci. Instrum.*, 1987, **20**, 1326-1330
- [10] Schwartz R., Precision Determination of Adsorption Layers on Stainless Steel Mass Standards by Mass Comparison and Ellipsometry. Part II: Sorption Phenomena in Vacuum, *Metrologia*, 1994, **31**, 129-136
- [11] Picard A. and Fang H., Methods to determine water vapour sorption on mass standards, *Metrologia* 2004, **41**, 333-339
- [12] Davidson S., Determination of the effect of transfer between vacuum and air on mass standards of platinum-iridium and stainless steel, *Metrologia*, 2010, **47**, 487-497
- [13] Berry J. et al, Report on the CCM WG TG1 pilot comparison to measure water vapour sorption on stainless steel mass standards, National Physical Laboratory Report ENG 46, 2013
- [14] Berry J. and Davidson S., Effect of pressure on the sorption correction to stainless steel, platinum/iridium and silicon mass artefacts, *Metrologia*, 2014, **51**, S107 – S113
- [15] Berry, J.; Borys, M.; Firlus, M.; Green, R.; Malengo, A.; Mecke, M.; Meury, P.-A.; Zůda, J.: Analysis of the correlation of sorption coefficients to pressure. National Physical Laboratory Report ENG 50, 2014
- [16] Chao L. S., Seifert F., Cao A., Haddad D., Newell D. B., Schlamminger S., and Pratt J., The design of the New NIST-4 watt balance, *CPEM Digest*, 2014, 364-5
- [17] Holmes F. T. and Beams J. W., *Nature*, 1937, **140**, 30
- [18] Holmes F. T., *Rev. Sci. Instrum.*, 1937, **8**, 444-7
- [19] Beams J. W., Hulburt C. W., Lotz W. E. Jr. and Montague R. M. Jr., *Rev. Sci. Instrum.*, 1955, **26**, 1181-5
- [20] Beams J. W. and Clarke A. M., *Rev. Sci. Instrum.*, 1962, **33**, 750-3
- [21] Gast Th., Waagen mit freischwebender magnetischer Aufhängung *Naturwissenschaften*, 1969, **56**, 434-8
- [22] Earnshaw S., *Trans. Cambridge Phil. Soc.*, 1839, **7**, 97–112
- [23] Berry M. V. and Geim A. K., *Eur. J. Phys.*, 1997, **18**, 307-13
- [24] Sinha P.K., *Electromagnetic suspension: dynamics and control* (London: Peter Peregrinus Ltd., 1987)
- [25] Stambaugh, C., The Control System for the Magnetic Suspension Comparator System for Vacuum-To-Air Mass Dissemination, *Proceedings of Asia-Pacific Symposium on Measurement of Mass, Force & Torque* (APMF 2015)
- [26] Stambaugh C., Mulhern E., Benck E., Kubarych Z., and Abbott, P., Progress on magnetic suspension for the NIST vacuum-to-air mass dissemination system, *XXI IMEKO World Congress "Measurement in Research and Industry"*, Prague, Czech Republic
- [27] Davidson S., Characterization of the long-term stability of mass standards stored in vacuum by weighing and surface analysis, *Metrologia*, 2012, **49**, 200–208
- [28] Lafferty J M *Foundations of Vacuum Science and Technology* (New York: Wiley-Interscience) 1998
- [29] Fuchs P., Low-pressure plasma cleaning of Au and PtIr noble metal surfaces, *Appl. Surf. Sci.*, 2009, **256** 1382–90
- [30] Marti K., Fuchs P. and Russi S., Traceability of mass II: a study of procedures and materials, *Metrologia*, 2015, **52**, 89-103

- [31] Fuchs P., Marti K. and Russi S., New instrument for the study of 'the kg, mise en pratique': first results on the correlation between the change in mass and surface chemical state, *Metrologia*, 2012, **49**, 607-614
- [32] Berry J., Downes S. and Davidson S., UV/ozone cleaning of platinum/iridium kilogram mass prototypes *Metrologia*, 2008, **47**, 410–418
- [33] ISO 4288:1996, "Geometrical Product Specifications (GPS) — Surface texture: Profile method — Rules and procedures for the assessment of surface texture," *ISO*, 1996
- [34] ISO 25178-2:2012, "Geometrical product specifications (GPS) -- Surface texture: Areal - Part 2: Terms, definitions and surface texture parameters," *ISO*, 2012
- [35] OIML R111-1 2004 (E), "Weights of classes E1, E2, F1, F2, M1, M1–2, M2, M2–3 and M3, Part 1: Metrological and technical requirements", *OIML International Recommendation*, 2004
- [36] R. Schwartz, "Precision Determination of Adsorption Layers in Stainless Steel Mass Standards by Mass Comparison and Ellipsometry Part1: Adsorption Isotherms in Air," *Metrologia*, **31**, 117-128, 1994
- [37] M. Kochsiek, "Measurement of Water Adsorption Layers in Metal Surfaces," *Metrologia*, **18**, 153-159, 1982

## Acknowledgements

The work leading to some of the results described in this publication is part of the European Metrology Research Programme (EMRP), which is jointly funded by the EMRP participating countries within the European Association of National Metrology Institutes (EURAMET) and the European Union.

Certain commercial equipment, instruments, or materials are identified in this paper in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the materials or equipment identified are necessarily the best available for the purpose.