

Referee report

Manuscript: *High-precision gravimetric primary standards for atmospheric Ar/N₂ measurements and redetermination of atmospheric Ar abundance (egusphere-2026-4405)*

Authors: N. Aoki, S. Ishidoya, N. Matsumoto

1. General comments and scope of this review

This manuscript describes an improved gravimetric preparation method for highly precise Ar standard mixtures (HPAs), in which CO₂, Ar, O₂ and N₂ are introduced independently from pure source gases rather than from a premixed CO₂/Ar mixture. The resulting standards achieve a standard uncertainty of approximately 5 per meg in $\delta(\text{Ar}/\text{N}_2)$, an improvement of more than an order of magnitude over the authors' previous results in 2019. This reduction in uncertainty opens the door to SI traceable (absolute) measurements of Ar/N₂ in atmosphere, which are currently anchored to several scales (sets of reference gas mixtures) with noticeable differences between them. Reference air mixtures with certified absolute values of Ar/N₂ associated with sufficiently low uncertainties would be a more comfortable framework for global and long-term atmospheric monitoring.

I have reviewed the scientific content of the manuscript, including the internal consistency of the reported values, the experimental methodology, the uncertainty budget, and the compatibility of the main conclusions with previously published determinations. I have not checked the language of the manuscript, which can be handled at the editorial level; the comments below are limited to scientific and technical issues, although a small number of numerical and bibliographic inconsistencies are included because they affect the scientific content.

2. Strengths of the manuscript

2.1 Rigor of the experimental method and measurements

The experimental work is of high quality and several aspects deserve explicit mention:

1. **Systematic diagnosis of the failed approach.** The initial preparation route based on premixed CO₂/Ar was abandoned after the authors identified systematic, lot-to-lot discrepancies of ~100 per meg in $\delta(\text{Ar}/\text{N}_2)$ — two orders of magnitude above the target signal. The candidate causes (weighing drift, residual gas in lines, transfer-induced fractionation) were evaluated one by one, and the correct one (thermal-diffusion fractionation during mixture transfer, consistent with the authors' own previous work) was identified. Moving to the independent introduction of pure CO₂, Ar, O₂ and N₂ elegantly removes the failure mode: a

pure gas cannot undergo compositional fractionation during transfer. This diagnostic sequence is a model of good metrological practice.

2. **High-precision weighing design.** Determining the transferred Ar mass (≈ 14 g) from the mass loss of a ~ 13 g cylinder, weighed on a 2.1 kg comparator with 0.01 mg sensitivity, rather than from the mass increase of the ~ 10 kg cylinder on a 26.1 kg comparator with 1 mg sensitivity, is the key precision gain (factor ~ 13 on the mass term). The weighing system was carefully characterized: reproducibility assessed by eleven cycle experiments with a slight cooling reproducing post-transfer temperature conditions, linearity verified with standard weights over 2–15 g.
3. **Complete treatment of surface and mechanical effects.** Pressure-dependent cylinder expansion was measured experimentally and the resulting buoyancy corrections (up to 14 mg, shifting the Ar molar fraction by $-0.08 \mu\text{mol mol}^{-1}$) are applied with uncertainties; permeation through the PCTFE valve seals was measured and shown to be negligible. These corrections are at the level of the claimed target uncertainty, and their experimental determination rather than estimation is valuable.
4. **Traceable impurity analysis.** Recognizing that traces of Ar in pure N_2 becomes a leading uncertainty contribution at the per meg level, the authors quantified it by GC–TCD against dynamically prepared standards traceable to an Ar/He primary mixture, with the Ar+O₂ co-elution handled by an independent galvanic-cell O₂ determination and an experimentally determined TCD sensitivity factor. The valve-switching procedure to purge the N_2 matrix is well justified by the baseline evidence shown.
5. **Isotopic control.** Determining the molar masses of Ar, O₂ and N_2 from measured isotopic abundances of the source gases rather than IUPAC values (representing natural variability) reduces this contribution by a factor of ~ 60 .
6. **Redundant analytical verification.** All ratios were measured with two independent mass spectrometers, and the eight HPAs were prepared over an 18-month period from different source gas cylinders, providing between-lot information on preparation reproducibility.

2.2 Quality of the uncertainty budget

The uncertainty budget follows the GUM and is, in my assessment, correctly constructed in its structure and reproducible in its numbers:

1. The measurement model (Eq. 9) and the law of propagation of uncertainty (Eq. 10) are stated explicitly, and Table 2 gives the grouped contributions (mass, molar mass, purity) corresponding to the three terms of Eq. 10.

2. The final uncertainty for the values in figure 10 were checked: propagating $u(\text{Ar}/\text{N}_2) = 1.3 \times 10^{-7}$, $u(\text{O}_2/\text{N}_2) = 1.8 \times 10^{-6}$ and $u(K) = 0.1 \mu\text{mol mol}^{-1}$ through Eq. (13) yields $0.102 \mu\text{mol mol}^{-1}$, in agreement with the stated $\pm 0.10 \mu\text{mol mol}^{-1}$.
3. The redetermined reference-air value (relative standard uncertainty 9.2 per meg) is consistent with the claimed factor-of-8 reduction ($0.9/0.11 = 8.2$), and the atmospheric value (10.9 per meg) is consistent with a quadrature combination of the reference value with a ~ 5.8 per meg measurement component.
4. The budget is transparent about which quantities dominate (transferred Ar mass and impurity Ar in pure N_2), which is valuable guidance for the community.

3. Points requiring clarification in the methods

Residual gas in the transfer lines. The residual fraction after the 100 (CO_2) and 200 (Ar) pressurization–expansion cycles is described as "estimated to be negligible relative to the transferred amount" based on the dilution factor of a single cycle, but no numerical estimate is given. Please provide it.

Permeation during weighing was tested for Ar only. The PCTFE permeation experiment was performed with pure Ar at 1 MPa. A brief justification (or test) for CO_2 , whose transfer mass (0.7 g) is 20 times smaller than that of Ar, would complete the assessment.

The new value sits above all previous determinations and this should be acknowledged explicitly. The reported atmospheric Ar molar fraction ($9335.71 \pm 0.10 \mu\text{mol mol}^{-1}$) is higher than every previously published value: Park et al. (2004) 9332 ± 3 , Sutour et al. (2007) 9330 ± 3.2 , Tohjima et al. (2005) 9333.2 ± 2.1 , and Aoki et al. (2019) $9334.4 \pm 0.7 \mu\text{mol mol}^{-1}$. The statement that "all estimates remain consistent within their expanded uncertainties" is formally correct. However, the only determination of comparable precision (Aoki et al., 2019) is from the same group and shares the EMRI/AIST scale anchor, so it provides limited independent verification. Also, taking the three fully independent determinations together (weighted mean $9332.2 \pm 1.5 \mu\text{mol mol}^{-1}$), the new value differs by $\sim 2.3 \sigma$. This may be explainable, but deserves an explicit discussion, including possible contributions (sampling year and location, scale differences, atmospheric variability).

Relevance for air-density metrology. The redetermined Ar abundance is $\sim 3.5 \mu\text{mol mol}^{-1}$ above the value $x(\text{Ar}) = 9332(3) \mu\text{mol mol}^{-1}$ underpinning the CIPM-2007 air-density formula. A brief note considering the possible impact of the result relative to that framework would be valuable for the metrology readership.

4. Clarifications required in the uncertainty budget

Numerical inconsistency for the redetermined reference value. The abstract and the conclusion give the redetermined absolute Ar/N_2 ratio of the AIST reference air as

$0.01195567 \pm 0.000000011$ (1.1×10^{-8}), while section 4.3.1 gives ± 0.00000011 (1.1×10^{-7}). Only the latter is consistent with the stated factor-of-8 uncertainty reduction ($0.9 \times 10^{-6} / 1.1 \times 10^{-7} = 8.2$). Please correct the abstract and conclusion.

Contradictory uncertainties for the CO₂ validation. Section 2.5.2 states that the CO₂ primary standards used to calibrate the Picarro G2301 have standard uncertainties of "approximately $0.80 \mu\text{mol mol}^{-1}$ ", whereas Table 5 reports measured CO₂ molar fractions with uncertainties of $0.05 \mu\text{mol mol}^{-1}$, which is questionable as it cannot be lower than the standards against which they are calibrated. Also, the two values imply very different validation strengths: with $u = 0.80 \mu\text{mol mol}^{-1}$, the $k = 2$ comparison envelope is $1.6 \mu\text{mol mol}^{-1}$ and the CO₂ agreement only weakly constrains transfer losses. With $u = 0.05 \mu\text{mol mol}^{-1}$, the envelope is $0.10 \mu\text{mol mol}^{-1}$ and the validation is powerful. Since the CO₂ agreement is the only external check on the transfer-and-weighing chain (no independent Ar/N₂ reference exists), please clarify this point, as it materially affects how strongly the CO₂ validation supports the Ar/N₂ assignments. Note that $0.80 \mu\text{mol mol}^{-1}$ would be implausibly large for modern SI-traceable CO₂ in air standards, but Table 1 displays uncertainties of $0.04 \mu\text{mol mol}^{-1}$, which appear very small. Please review and explain the apparent inconsistencies.

Covariance treatment. Equation (10) is a root-sum-square without covariance terms. However, the molar fractions $y(\text{Ar})$, $y(\text{O}_2)$ and $y(\text{N}_2)$ share the common denominator Σn and are therefore strongly correlated (dominated by the O₂ filling-mass uncertainty). Please state if correlations were also considered or explain why not. Likewise, the 8 HPAs share source-gas cylinders (one pure Ar cylinder, three to four N₂ cylinders), so their gravimetric values are correlated, but the least-squares regression treats them as independent. Given that this manuscript itself demonstrates (Fig. 5) that lot-level systematic effects are the dominant hazard in this field, a comment on between-lot correlations is warranted.

Breakdown of $u(K)$. The sum $K = 999\,552.8 \pm 0.1 \mu\text{mol mol}^{-1}$ (Section 4.3.2) is decomposed with all values of the trace gases provided, but without uncertainty. Please provide the uncertainties for each compound.

Table 2 rounding. The stated combined values for N₂ ($0.76 \mu\text{mol mol}^{-1}$) and CO₂ ($0.032 \mu\text{mol mol}^{-1}$) are not reproduced by the RSS of the displayed components, presumably because of rounding of the individual contributions. One additional digit would make the table verifiable.

5. Comments on vocabulary and terms

Molar fraction. The "molar fraction" is not the recommended term for the quantities in which an amount-of-substance of the analyte is divided by the amount-of-substance of the matrix gas. The CIPM and CCQM now recommends the term "amount-of-substance fraction", which may be shortened in "amount fraction". Consider changing or introduce

the terminology in the beginning of the article, explaining why “molar fraction” was chosen.

Expressing numbers with many digits. In view of the very low uncertainty reported on numbers with many digits, consider expressing the numbers with a format allowing easy reading, for example grouping the digits by three (see the SI Brochure on bipm.org).

6. Comments on figures and tables

Table 2 needs reformatting to separate the inner table listing contributions from the main table displaying the uncertainty sources and the combined uncertainty for each component of the gas mixture. The inner table on contributions is difficult to understand and needs clarification.

Table 5 suffers from a formatting issue, with the document line numbers appearing on top of the extreme right column values. In the same table, two rows are displayed for each cylinder for two values obtained with different instruments. It is valuable but the formatting could be improved, for example by using light colors for the background of all values for one instrument.

7. References to be added or corrected

Missing reference Sutour, C., Stumpf, C., Kosinski, J.-P., Surget, A., Hervouët, G., Yardin, C., Madec, T., and Gosset, A.: Determination of the argon concentration in ambient dry air for the calculation of air density, *Metrologia*, 44, 448–452, <https://doi.org/10.1088/0026-1394/44/6/003>, 2007. This is one of the few absolute determinations of atmospheric Ar ($9330 \pm 3.2 \mu\text{mol mol}^{-1}$) and, together with Park et al. (2004), the basis of the $x(\text{Ar})$ value adopted in the CIPM-2007 air-density formula. It should be included in Fig. 10 and in the comparison in section 4.3.2.

Missing reference The CIPM-2007 air-density formula (Picard, A., Davis, R. S., Gläser, M., and Fujii, K.: Revised formula for the density of moist air, *Metrologia*, 45, 149–155, 2008) is relevant context as well.

Missing reference "Birner et al. (2020)" is cited in section 3.1 (line 235) but does not appear in the reference list. Please add the full citation.

Missing reference: "WMO (2021)" is cited in section 4.3.2 for the CH₄ and N₂O global annual means but is absent from the reference list. Please add the full citation (e.g., WMO Greenhouse Gas Bulletin No. 17, 2021, or the specific data source used).

Incorrect DOI: the entry for Ishidoya et al. (2021) cites the journal as "Atmos. Chem. Phys., 21, 1357–1373" but gives the DOI <https://doi.org/10.5194/amt-14-6181-2021>, which belongs to a different paper (Aoki et al., *Atmos. Meas. Tech.*, 14, 6181–6193, 2021). The correct DOI for the Ishidoya et al. paper is <https://doi.org/10.5194/acp-21-1357-2021>.

Author name misspelling: "Böhlk" should be "Böhlke" (Böhlke, J. K.: Variation in the terrestrial isotopic composition and atomic weight of argon, *Pure Appl. Chem.*, 86, 1421–1432, 2014). The misspelling occurs in sections 2.4 and 4.1.2 and in the reference list.

6. Recommendation

The manuscript represents a sincere metrological advance — the first Ar/N₂ primary standards with uncertainties at the level of the analytical systems they are meant to calibrate, and the main conclusions are compatible with the published literature. None of the points raised requires new measurements. I therefore recommend publication subject to minor-to-moderate revisions addressing the points above, and I would be glad to review the revised manuscript.