Supplemental Material: Revision of the WMO/GAW CO_2 Calibration Scale

1 Notes on historical manometric CO_2 determinations and corrections applied

1.1 Historical manometric records

Manometric CO_2 mole fractions from 1996-2017 were determined based on the maximum CO_2 mole fraction calculated after the gas in the small volume was allowed to warm to oven temperature. Using a database of manometric results as reference, we examined raw data files and flagged data records when a) the CO_2 maximum could not be clearly identified (not common), b) the data record was not sufficient to determine a loss rate (e.g. too little data), c) raw files corresponding to a result in our database could not be found, or d) data appeared abnormal, e.g. temperatures not converging or showing abnormal variability. We were able to recover and apply corrections to 93% of the original data records.

Most manometer runs show one maximum in the calculated CO_2 record. However, some manometer runs from 1998 and some from 2004 show two CO_2 peaks (Fig. S1). This secondary peak could be related to H_2O desorbing from surfaces in the small volume. In these cases, we calculated the loss correction using the time associated with the first CO_2 peak, and the slope from the section of data after the second peak.

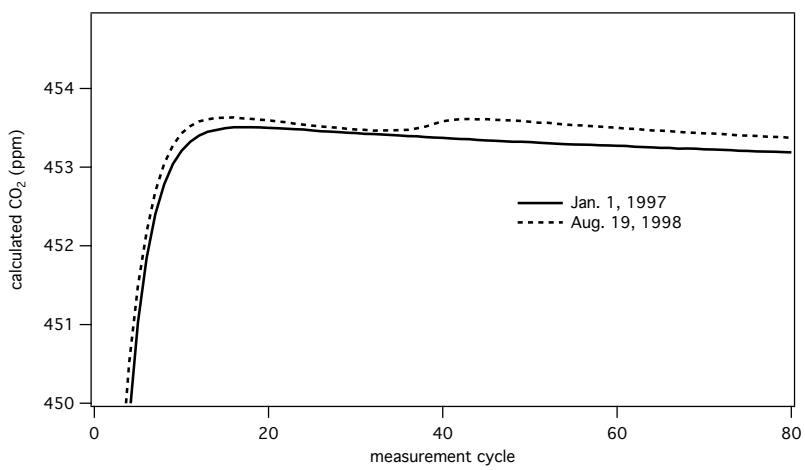


Figure S1: CO_2 mole fraction calculated from the pressure of gas in the small volume and oven temperature measured at various locations within the oven, as function of time (measurement cycle), where each cycle is ~60s. During the final extraction step, the small volume is cooled using a small dewar of liquid nitrogen. The dewar is removed after all CO_2 has been collected in the small volume. As CO_2 sublimates and equilibrates to the oven temperature, the calculated CO_2 reaches a maximum 15-20 minutes (~16 measurement cycles in this case) after the L-N_2 was removed. Shown are manometer data from two different runs of primary standard AL47-107. The January 1997 run shows the typical CO_2 maximum, followed by a slow decline. The August 1998 run shows a second peak around measurement cycle 40. For consistency, we applied corrections to existing CO_2 data based on timing of the first CO_2 peak.

1.2 CO_2 Loss Rate

We made corrections for CO_2 loss assuming a linear loss rate. Extended data records indicate that the loss rate decreases with time (Fig. S2). However, in the short term, the initial loss rate is approximately constant and can be approximated by a linear function (Fig. S2 inset).

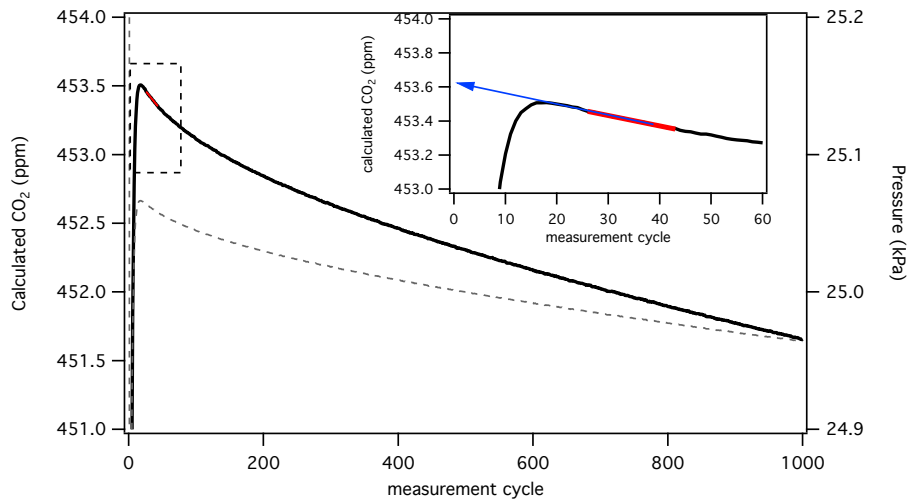


Figure S2: Extended manometric run showing the non-linear decline of CO_2 in the small volume (solid line) and the pressure in the small volume (dashed line) (temperature is nearly constant after cycle 20). The calculated CO_2 loss rate is higher at first, and then gradually slows. The inset shows a close-up view of the first 60 measurement cycles (~30 s per cycle). While the true CO_2 loss rate is non-linear, the loss rate is approximately linear in the vicinity of the maximum CO_2.

Fig. S3 shows the results of a pure CO_2 injection test, in which pure CO_2 was introduced into the small volume through an auxiliary port, without the cryogenic extraction step. The goal was to observe CO_2 loss upon immediate injection of CO_2 (without a large temperature change). This period is not observable in most data files since the small volume temperatures and pressure are changing rapidly during the initial minutes. Fig. S3 suggests that the loss of CO_2 begins immediately, and is non-linear. Linear extrapolation from around the time of the typical CO_2 peak to the beginning of the record (~100 s) results in a CO_2 result that is ~0.2 ppm lower than observed from the CO_2 injection experiment. However, the conditions of this experiment are not the same as those during an extraction procedure. With direct injection, surfaces, including o-rings, in the manometer are exposed to higher CO_2 pressures and temperatures compared to when CO_2 is extracted from air. Thus, we consider the 0.2 ppm underestimate suggested from these data (Fig. S3) an upper limit. These experiments prompted the preparation of gravimetric CO_2-in-air standards as way to independently assess the magnitude of the proposed scale correction.

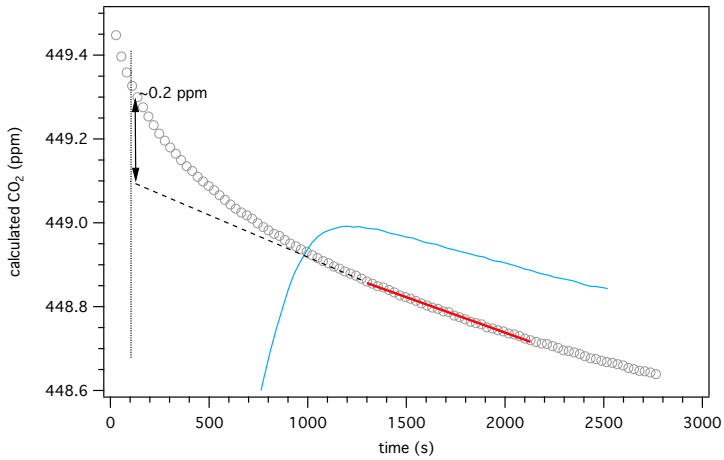
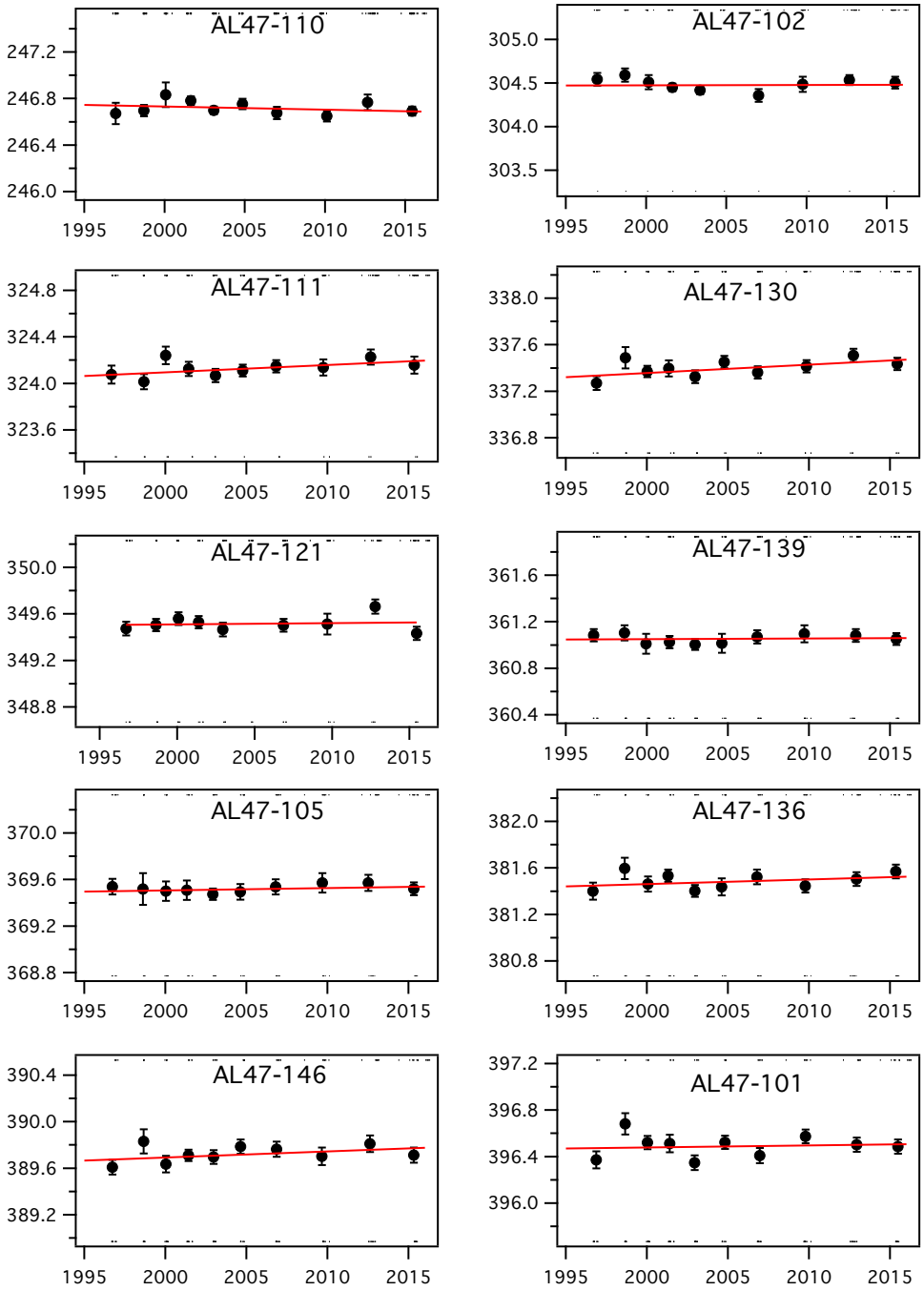


Figure S3: Test in which pure CO_2 was introduced into the small volume directly, without a cryogenic extraction step. Open symbols denote the CO_2 mole fraction that would correspond to this sample. The red line is a linear fit to the open symbols. The blue line is a representative extraction of a 450 ppm CO_2 sample, in which the purified CO_2 starts cold, and warms over time, showing the characteristic CO_2 maximum.

1.3 Drift Assessment


AL47-110
247.2
246.8
246.4
246.0
AL47-102
305.0
304.5
304.0
303.5
AL47-111
324.8
324.4
324.0
323.6
AL47-130
338.0
337.6
337.2
336.8
AL47-121
350.0
349.6
349.2
348.8
AL47-139
361.6
361.2
360.8
360.4
AL47-105
370.0
369.6
369.2
368.8
AL47-136
382.0
381.6
381.2
380.8
AL47-146
390.4
390.0
389.6
389.2
AL47-101
397.2
396.8
396.4
396.0
1995
2000
2005
2010
2015

60

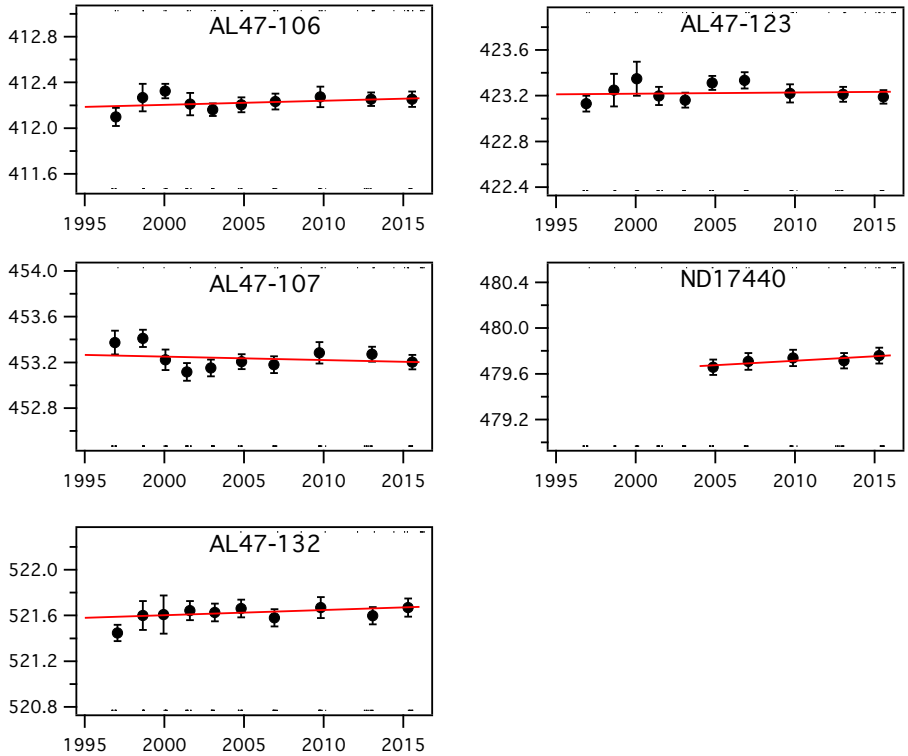


Figure S4: Average manometric results for each episode (1-sigma error bars), and associated weighted linear fits versus time (solid lines).

Table S1: Average mole fractions and drift rates (ppm decade^{-1}) for 15 primary standards with measurement histories spanning at least 10 years. Drift rates were determined from a weighted linear fit to manometer data, averaged by episode, over N_{ep} episodes. Uncertainties are 95% C.L.

Cylinder	Avg. (x2019)	Drift Rate	Uncertainty	N_{ep}
AL47-110	246.724	-0.0268	0.0624	10
AL47-102	304.495	0.0030	0.0855	9
AL47-111	324.134	0.0628	0.0879	10
AL47-130	337.403	0.0717	0.0719	10
AL47-121	349.515	0.0099	0.0747	9
AL47-139	361.054	0.0061	0.0699	10
AL47-105	369.523	0.0208	0.0854	10
AL47-136	381.487	0.0393	0.0792	10
AL47-146	389.731	0.0530	0.0847	10
AL47-101	396.495	0.0175	0.0808	10
AL47-106	412.231	0.0365	0.0898	10
AL47-123	423.218	0.0104	0.0904	10
AL47-107	453.255	-0.0303	0.0922	10
ND17440	479.720	0.0772	0.2573	5
AL47-132	521.605	0.0694	0.1042	10

2 Uncertainty Analysis

2.1 Temperature Measurement during CO_2 determination

Standard uncertainties reported by accredited laboratories for calibration of three platinum resistance thermometers (PRTs) are 0.0125 °C. The device used to measure resistance (Hart 1529) adds an additional 0.007 °C. Combining these in quadrature we get 0.014 °C for a single probe. We use the average temperature determined from 2-3 probes placed in the vicinity of the large and small volumes to determine T. Combining the uncertainty from 3 probes with the typical standard deviation observed during the measurement we obtain:

$$uT_{air} = \sqrt{(\frac{0.014}{\sqrt{3}})^2 + 0.004^2} = 0.009\ °\mathrm{C} \quad (s1)$$

$$uT_{co2} = \sqrt{(\frac{0.014}{\sqrt{2}})^2 + 0.014^2} = 0.017\ °\mathrm{C} \quad (s2)$$

uT_{CO2} is larger than uT_{air} because the oven must respond to a change in temperature (oven door open briefly for cryogenic trapping of CO_2 in the small volume).

2.2 Pressure Measurement during CO_2 determination

Pressure is measured using a Paroscientific 6000-15A pressure sensor. The 6000-15A is calibrated in-house using a Ruska 2465 piston gauge. We apply a linear calibration (span, offset) to the pressure output (Fig S5). Secondary checks on the stability of the span have been determined relative to a secondary pressure standard (Fluke, RPM4) at 10 kPa increments from 10-90 kPa since 2015. Since 2008, the span has varied between 0.99977 and 0.99999. The offset is determined by reading the pressure at vacuum following initial evacuation of the manometer. The offset recorded at vacuum usually agrees within ~0.003 kPa with that determined from the intercept using a linear fit to the piston gauge data. From 2004-2012 the offset varied between 0.001-0.005 kPa. A larger offset has been observed (0.018-0.024 kPa) since 2014 following a calibration by the manufacturer.

The estimated accuracy of the piston gauge is 5 parts-per-million (ppm) at full scale (100 kPa), 6-8 ppm at mid-scale, and 13 ppm at 10 kPa (manufacture's specifications). In addition, we observe drift in the 6000-15A "offset", or zero, of 0.0005 kPa over a few hours. The repeatability of the pressure measurement over short time periods (a few minutes) is excellent, often <0.0002 kPa, 1-sigma. Our estimates of uP are mainly functions of the accuracy of the pressure calibration and stability of the device over the course of an extraction (3 h). It is also possible that small leaks or outgassing in the system would lead us to over-estimate P_{CO2} since P_{CO2} is always less than room pressure and the measurement of P_{CO2} takes ~20 minutes to complete. System leaks are not expected to contribute significantly to other pressure measurements because they are made without significant lag-time. Recent

experiments performed at vacuum and also with 30 kPa air in the small volume indicate out-gassing and/or leaks that induce pressure changes as high as 4 Pa hr^{-1} in the absence of CO_2. A 4 Pa hr^{-1} leak would lead to ~0.02 ppm overestimate of X_{CO2}. Since we do not have similar data for all episodes, we include the potential for leaks in the uncertainty associated with P_{CO2}. Combined estimates of uncertainty in pressure are as follows:

For P_{air}, typically 83 kPa:

uP_{air} components: piston gauge ($5 \cdot 10^{-6} * P_{air}$), zero drift (0.0002 kPa), repeatability (0.0001 kPa)

uP_{air} = 0.0007 kPa

For P_{CO2}, at 30 kPa:

uP_{CO2} components: piston gauge ($8 \cdot 10^{-6} * P_{CO2}$), zero drift (0.0004 kPa), leak (0.0003 kPa), repeatability (0.0001 kPa)

uP_{CO2} = 0.0010 kPa

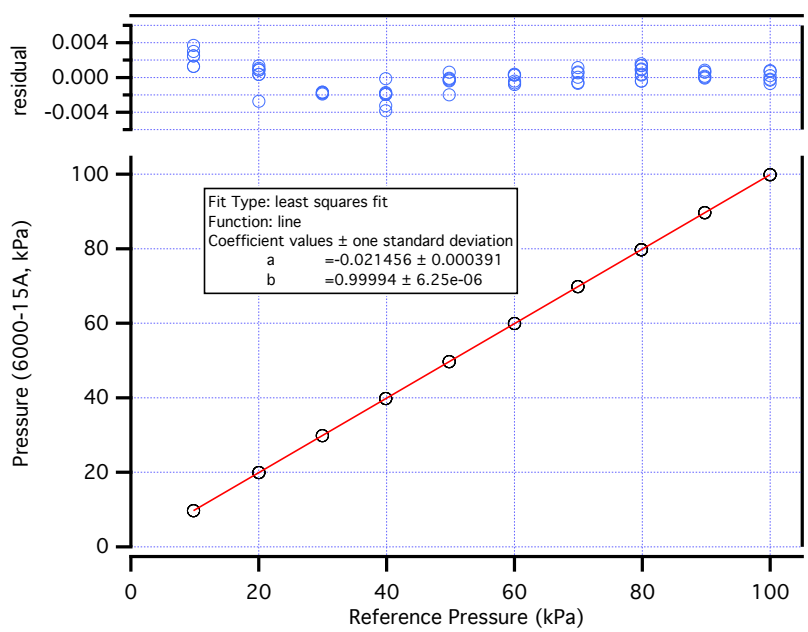


Figure S5: Example calibration of the main pressure sensor using the piston gauge, from 2014.

2.3 Volume Ratio

The volume ratio, Φ, is calculated from a series of four gas expansions, using two auxiliary glass volumes. For each expansion, we calculate an intermediate volume ratio, r_i. Each intermediate volume ratio is determined from measurements of temperature, pressure, and second virial coefficients for the expanding gas (usually air or nitrogen) (Zhao *et al.*, 1997).

The overall volume ratio is calculated as:

$$\Phi = r_1 * r_2 * r_3 * r_4 - r_1 * r_2 * r_3 + r_1 \qquad \text{(s3)}$$

The repeatability of volume ratio experiments is often very good. For example, for 24 volume ratio experiments run over 6 days in January 2017, the mean and standard deviation were 880.10 and 0.05 respectively. This repeatability is similar to that described in (Zhao *et al.*, 1997). However, we need to consider uncertainties in T and P measurements, and compare Φ calculated using different gases in order to estimate the uncertainty (Figure S6b).

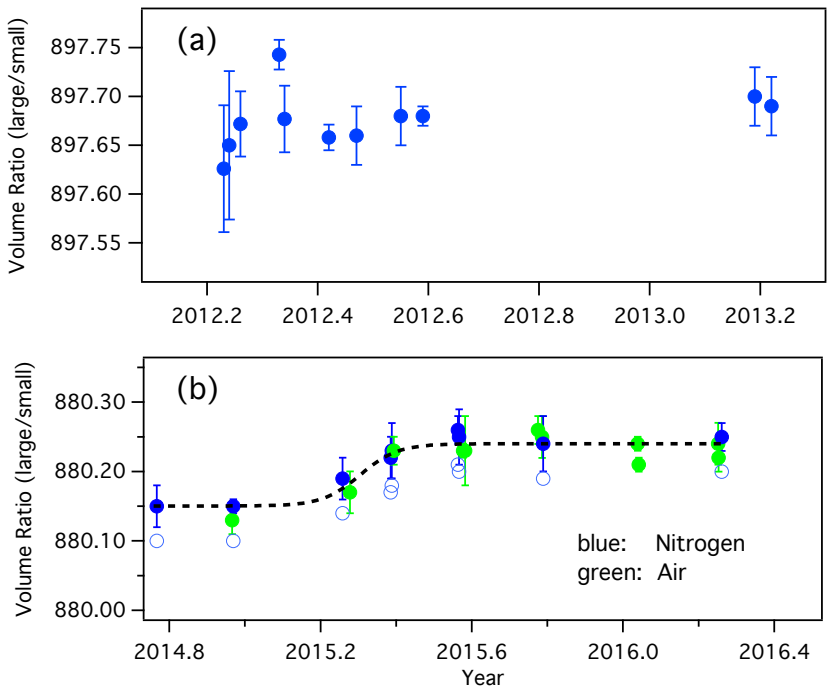


Figure S6: Results from volume ratio expansion experiments (a) 2012-2013 and (b) 2014-2016. Shown are the mean and standard deviation from four experiments performed in a given day, using different gases: CO_2-free air (green), and nitrogen (blue). Open circles in (b) show volume ratios calculated using an alternate second virial coefficient for N_2: -4.64 cm^3/mol at 300 K (Dymond *et al.*, 2002) instead of -5.18 cm^3/mol (https://web.archive.org/web/20190506031327/http://www.kayelaby.npl.co.uk/). Note that the small volume was replaced in 2014, so the volume ratios in (a) and (b) are not expected to be the same. Volume ratios performed in 2019 using argon gas were found, on average, to be 0.018% lower than those shown in (b).

2.3.1 Effect of uncertainties in temperature on volume ratio

We begin with the estimate for uT_{air} from 2.1 since the oven is not disturbed during a volume ratio experiment. On each expansion, we measure temperature as the average of two PRT probes placed in a central location in the oven, with an initial estimate for uT = 0.009 °C, as before. However, during pressurization and expansion steps we observe adiabatic heating and cooling, which could lead to temperature gradients in the glass volumes. During a normal volume ratio experiment, we allow seven minutes for equilibration prior to taking measurements. By placing

4-6 calibrated thermistors at various locations on each volume (22 in total), we observed variations in temperature (σ=0.03 °C, 6 probes). Variations were largest during volume expansions 3 and 4 (corresponding to r_3, r_4), and negligible on expansions r_1 and r_2. Therefore, we include additional uncertainty for r_3 and r_4.

$$uT\ (r_1\ \&\ r_2) = 0.01\ °C \tag{s4}$$

$$uT\ (r_3\ \&\ r_4) = sqrt((0.009^2 + (0.03/sqrt(6))^2) = 0.015\ °C \tag{s5}$$

2.3.2 Effect of uncertainties in pressure on volume ratio

For pressure, we apply the same logic as in 2.2, except that we increase contributions from potentials leaks and zero drift because larger volumes are held at vacuum for ~1 hour on each experiment. We assume zero drift uncertainties of 0.0002 kPa, and leak contributions ranging from 0-0.0005 kPa. Repeatability is a small component (0.0001 kPa). Listed here are the components corresponding to uncertainties on each pressure measurement (manufacturer's specification, zero drift, leak potential).

On filling to 80 kPa,	$5{\cdot}10^{-6}$*80, 0.0002, 0.0000;	uP = 0.0006 kPa.
On expansion to 19 kPa,	$8{\cdot}10^{-6}$*19, 0.0002, 0.0004;	uP = 0.0006 kPa
On expansion to 10 kPa,	$13{\cdot}10^{-6}$*10, 0.0002, 0.0004;	uP = 0.0007 kPa
On expansion to 18 kPa,	$8{\cdot}10^{-6}$*18, 0.0002, 0.0005;	uP = 0.0008 kPa
On expansion to 13 kPa,	$12{\cdot}10^{-6}$*13, 0.0002, 0.0005;	uP = 0.0007 kPa

2.3.3 Second virial coefficients

Uncertainty in the second virial coefficient has only a small impact on the calculated volume ratio (Figure S6b). Using two different estimates of the second virial coefficient for nitrogen results in a 0.05 difference (0.0056%) in the volume ratio. Assuming a uniform distribution, this would add additional (0.05/2/sqrt(3))=0.014 uncertainty to the estimate of the volume ratio, Φ. Instead, we include an uncertainty component derived from testing with different gases (see 2.3.4).

2.3.4 Propagation of uncertainty for the Volume Ratio

Estimated uncertainties for each expansion step are shown in Table S2. We made some assumptions about the correlations of uncertainties since the same pressure transducer and temperature sensors are used for each measurement of P and T. For pressure, we assume high correlation between P_1 and P_2 (r=0.9). For T_1 and T_2, we assume a correlation coefficient of 0.5, since some of the uncertainty is related to adiabatic compression/expansion and would be uncorrelated.

Table S2: Parameters of a typical series of volume expansions, each from P_1 ==> P_2, used to determine the volume ratio.

P_1 (kPa)	uP_1 (kPa)	P_2 (kPa)	uP_2 (kPa)	r_i	ur_i
80	0.0006	19	0.0006	4.15	0.00017
80	0.0006	10	0.0007	7.95	0.00057
80	0.0006	18	0.0008	4.53	0.00025
80	0.0006	13	0.0007	6.85	0.00039

To estimate the total uncertainty in the volume ratio, we used the NIST metRology package (https://www.nist.gov/programs-projects/metrology-software-project) to evaluate equation s3 using parameters and uncertainties in Table S2. The uncertainty calculated from parameters in Table S2 is 0.105 for a volume ratio of 880. We add to that uncertainty contributions from temperature probe placement (0.08), differences between volume ratios determined using different gases (air, nitrogen, and argon) (0.03), and repeatability of volume ratio experiments (0.032), summed in quadrature: $u\Phi$ = sqrt(0.105^2+0.08^2+0.03^2+0.032^2). In total, we arrive at a standard uncertainty of $u\Phi$ =0.139 for a volume ratio of 880.1, or 0.016%. This uncertainty estimate is about 50% larger than that reported by (Zhao and Tans, 2006). Since the physical configuration of the NOAA manometer has remained basically the same since 1996, although with some changes related to valves and pressure sensors, we assume that the relative uncertainty of the volume ratio is applicable to all volume ratios used since 1996.

2.3.5 Water vapor

While we do not have measurements of the water vapor content in the extracted CO_2, we can estimate X_{H2O} from the temperature of the traps. Both trap 1 and trap 2 are typically held at -67 °C during the separation step. The vapor pressure of H_2O over ice at -67 °C is 0.406 Pa (Huang, 2018). The estimated volume of trap 1 is 40 cm^3. For a 400 ppm CO_2 sample, 0.1 mmoles of CO_2 would be present in the trap, corresponding to a pressure of 6.5 kPa. The fraction of gas in trap 1 that is water vapor is then 0.406/(6500+0.406) =$6.25\cdot10^{-5}$. As this gas is trapped in the small volume and then warmed, water vapor would exert a pressure equivalent to ($6.25\cdot10^{-5}$)(400) = 0.025 ppm CO_2. There is no active control on the trap temperatures (we use an alcohol bath), and the typical range is -65 °C to -70 °C. For this range of temperatures, water vapor would correspond to 0.033 to 0.016 ppm CO_2, respectively. We therefore include a conservative uncertainty of 0.03 ppm due to water vapor (Table S3).

2.4 Total uncertainty associated with the manometric measurement

To arrive at the total uncertainty associated with a single manometric measurement of CO_2, we use the NIST metRology package and equation (s8),

$$X_{CO2} = (\phi^{-1})\frac{P_{CO2}T_{air}}{P_{air}T_{CO2}}(1 + A_1 - A_2) - X_{N2O} + X_{virial_{correction}} + X_{loss_{correction}} - X_{H2O} \quad \text{(s6)}$$

$$A_1 = \frac{P_{air}\beta_{air}}{RT_{air}}$$

$$A_2 = \frac{P_{CO2}\beta_{CO2}}{RT_{CO2}}$$

where T and P are the temperatures and pressures of the large volume (air) and small volume (CO_2), β_{air} and β_{CO2} are second virial coefficients, R is the gas constant, Φ is the volume ratio, and X_{N2O} is the mole fraction of N_2O in the air sample (see main document). We do not calculate an uncertainty for each individual measurement. Instead we use typical values and uncertainties for variables in (s8), and calculate uX_{CO2} over a range of mole fractions. This is justified because the measurement conditions do not vary significantly between measurements.

For each measurement in the database, we calculate a loss correction and a virial correction. The uncertainty associated with $X_{virial_correction}$ is ~0.005 ppm. For $X_{loss_correction}$, we estimate the uncertainty in loss rate at 15% for most measurements, and 25% for those exhibiting a second maxima. We assume that the time corresponding to peak CO_2 (t) is known to within one measurement cycle, and that the initial time (t_o) has an uncertainty of 2 minutes (2-4 measurement cycles). Together, the uncertainty associated with the loss correction is ~15%.

Propagating uncertainties in Table S3 through equation s8, we estimate the uncertainty associated with a single manometric measurement to be ~ 0.079 ppm at 400 ppm, and ranges from 0.06 ppm at 250 ppm to 0.11 ppm at 600 ppm (Table S4). Uncertainty in the volume ratio is the largest component, accounting for nearly half.

Table S3: Typical values and uncertainties for variables associated with measurement of a 400 ppm air sample.

Variable	Typical value	Standard uncertainty	unit	Approx. Relative contribution (%)
P_{air}	83	0.0007	kPa	1
P_{CO2}	29	0.001	kPa	3
T_{air}	310	0.01	K	2
T_{CO2}	310	0.017	K	6
Φ (vol. ratio)	880.10	0.139	dimensionless	48
B_{air}	-5.87	0.2	cm^3/mol	< 1
B_{CO2}	-112.8	0.2	cm^3/mol	< 1
$X_{virial_correction}$	0.03	0.005	ppm	1
$X_{loss_correction}$	0.14	0.02	ppm	6
X_{N2O}	0.325	0.0005	ppm	*negligible*
X_{H2O} (note 1)	0	0.03	ppm	11
^{13}C, ^{18}O	-8.5, -2	0.3	per mil	1 (note 2)
Repeatability (manometer)		0.039	ppm	19

Reproducibility, (scale transfer)		0.01	ppm	2

[1] We do not make a correction for water vapor, so its value is assumed to be zero.

[2] Estimated to be less than 0.01 ppm.

2.5 Total uncertainty, including scale transfer

The manometer repeatability from up to 10 episodes is relatively consistent over the range 250-800 ppm, with an average of 0.10 ppm. Since critical components are calibrated to national standards prior to each episode (temperature, length, mass) we consider episodes as independent. Thus, we calculate the manometric repeatability as the average standard error, from N episodes, among all primary standards. The average standard error for all primaries is 0.039 ppm.

Scale transfer from primary to secondary and secondary to tertiary standards is performed using multiple laser spectroscopic techniques (Tans *et al.*, 2017). Differences among tertiary cylinders analyzed more than once over different time periods, spread across multiple suites of secondary standards and multiple primary-secondary comparisons, captures the reproducibility of the scale transfer process. We estimate this uncertainty contribution to be 0.01 ppm (1σ).

Adding uncertainties associated with manometer repeatability and scale transfer in quadrature to the manometer uncertainty we obtain the total standard uncertainty associated with the X2019 scale (third column in Table S4). For expanded uncertainty (~95% Confidence Level) we multiply standard uncertainties by coverage factor k=2. The expanded uncertainty can be approximated as $uX_{CO2} = 0.000287*X_{CO2} + 0.074$ ppm.

Table S4: Standard and expanded uncertainties (ppm) associated with X_{CO2} determination (scale X2019).

X_{CO2} (µmol/mol)	Standard Uncertainty (manometer only)	Standard Uncertainty (including manometer repeatability and scale transfer)	Expanded Uncertainty (k=2)
250	0.056	0.074	0.148
300	0.063	0.080	0.160
350	0.071	0.087	0.173
400	0.079	0.093	0.187
450	0.087	0.100	0.201
500	0.095	0.108	0.216
550	0.103	0.116	0.232
600	0.112	0.124	0.248

3 Comparing primary standards (with manometric CO_2 determination) to gravimetrically-prepared standards

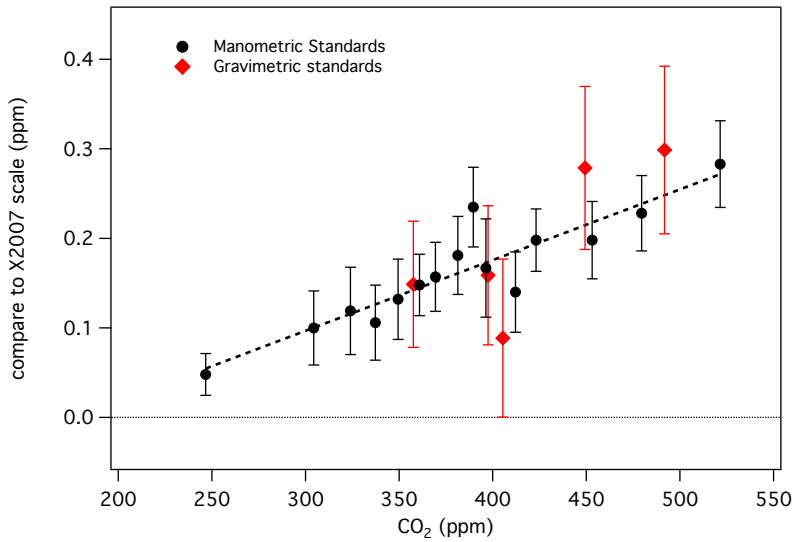


Figure S7: Differences between the X2019 scale and the X2007 scale derived from manometric standards (black symbols, dashed line), and differences between gravimetric standards and the X2007 scale (red symbols). Here, gravimetric standards were analyzed by laser spectroscopy and assigned CO_2 mole fractions on the X2007 scale. Error bars (1σ) are larger for gravimetric standards because they are independent of the manometric system and incorporate uncertainties in both methods, whereas manometric primary standards share common uncertainty components. The gravimetric standards show a mole-fraction dependent offset from X2007 similar to X2019 primary standards, suggesting that the linear correction for CO_2 loss in the manometer small volume is a reasonable approach.

References

Dymond, J. H., K. N. Marsh, R. C. Wilhoit, and K. C. Wong (2002), *Virial Coefficients of Pure Gases and Mixtures, Subvolume A: Virial Coefficients of Pure Gases, Numerical Data and Functional Relationships in Science and Technology*, Springer-Verlag, New York.

Huang, J. (2018), A Simple Accurate Formula for Calculating Saturation Vapor Pressure of Water and Ice, *Journal of Applied Meteorology and Climatology*, *57*(6), 1265-1272.

Tans, P. P., A. M. Crotwell, and K. W. Thoning (2017), Abundances of isotopologues and calibration of CO2 greenhouse gas measurements, *Atmos. Meas. Tech.*, *10*(7), 2669-2685.

Zhao, C. L., P. P. Tans, and K. W. Thoning (1997), A high precisionmanometric systemfor absolutecalibrations of CO2 in dry air, *J. Geophys. Res.*, *102*(D5), 5885-5894.

Zhao, C. L., and P. P. Tans (2006), Estimating uncertainty of the WMO mole fraction scale for carbon dioxide in air, *J. GeoPhys. Res.*, *111*.