G2307: Traceable calibration of Formaldehyde (H_2CO)

Jonathan Bent, Ph.D., Picarro Senior Application Scientist
Colin Wallace, Ph.D., Picarro Systems Engineer
Gregor Lucic, Ph.D., Product Manager
Chris Rella, Ph.D., Picarro Research Fellow
John Hoffnagle, Ph.D., Principal Spectroscopist
Karsten Baumann, Dr.-Ing, Application Scientist

ABSTRACT

In this white paper, we discuss the rescaling of, and establishment of a traceable calibration for, formaldehyde (H_2CO) on Picarro's G2307 "LBDS" Cavity Ring Down Spectrometers.

VERSION HISTORY

Version	Release Date	Authors	Notes
0.1	2022 06 08	Bent	
1.0	2023 03 27	Bent	

CONTENTS

INTRODUCTION

Picarro first started working on a formaldehyde instrument in ~2008, with the external and internal designations of G1107 and LADS, respectively. The instrument relied on a near-infrared spectral feature near 1514 nm. This instrument was adapted to the G2000 platform in ~2013 with the name G2107 (LADS) and used primarily by customers monitoring emissions from industrial facilities. In ~2016, Picarro undertook an effort to enable measurements of Semiconductor-relevant species at longer wavelengths where certain diatomic species absorb more strongly. Because of new hardware developed during this process, including mirrors and photodetectors, Picarro was able to successfully measure a formaldehyde peak near 1777 nm, providing roughly a factor of 100 improvement in sensitivity. Picarro redesigned the LADS into the G2307 (LBDS) instrument, and with this new laser was able to equip the instrument with a key new feature: surrogate gas validation, which allowed the health of the instrument to be checked by a safer, WMO-traceable gas: methane (CH_4). Importantly, this surrogate gas method allowed the LBDS instrument's overall drift and performance to be well characterized without the customer having to rely on formaldehyde reference materials which are hazardous, and can degrade with time.

SPECTROSCOPY AND SCALING

Spectroscopy

Picarro measures Formaldehyde (H_2CO, also HCHO, CH_2O) on the G2307 (LBDS) with a narrow-band tunable diode laser which produces light frequencies in the near-infrared region of the electromagnetic spectrum. Figure 1 shows the spectral region that the instrument uses to quantify H_2CO, H_2O, and CH_4 in units of inverse centimeters (cm^{-1}, Y [cm^{-1}] = 10,000,000 / X [nm]).

The CRDS instrument rapidly (<2 seconds) scans this spectral region. The resulting spectrograms are analyzed using a non-linear least squares optimization algorithm, using pre-calculated model functions for each of the spectral features in this region. The concentrations of H_2CO, H_2O, and CH_4 are reported by the instrument on the user interfaces and the data logs.

The CH_4 channel on the G2307 allows the user to characterize the absorption response of the analyzer without the need for the primary gas.[1] A successful linear response of the CH_4 channel requires that the

---

[1] Formaldehyde absorbs about 3.3x more strongly per ppm of gas than methane when considering the absorbance of the major features in this range.

components of the instrument that could be subject to drift—namely pressure and temperature sensors, laser, and photodetector—respond in an accurate and linear fashion. For a full explanation of the purpose of CH_4 as a surrogate gas, refer to Bent (2022).

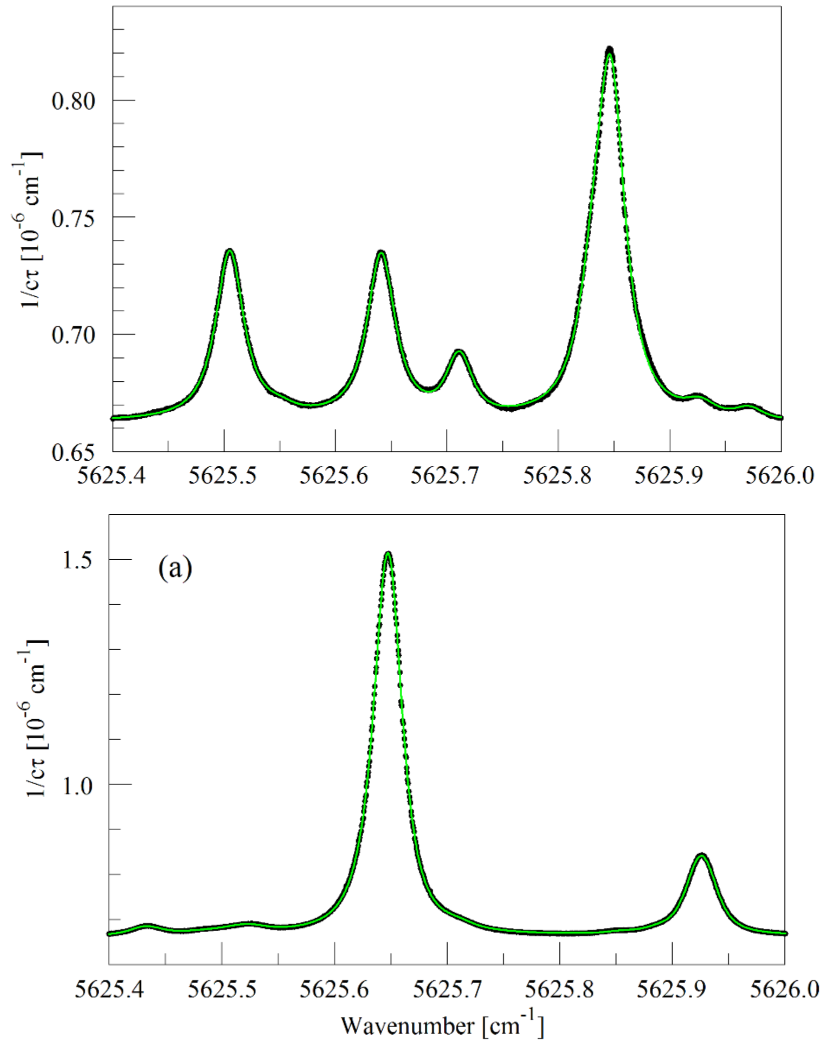


Figure 1. [Top] H_2CO spectral region, showing major features for H_2CO at 7.4 ppm concentration, including the primary peak at 5625.84 wavenumbers. A smaller feature of water can be seen at 350 ppm at 5625.71 wavenumbers. [Bottom] A spectrogram of CH_4 at 100 ppm across the same wavenumber range with major peaks at 5625.65 and 5625.92 wavenumbers. Methane can be used as a "surrogate" to check the calibration and proper operation of the instrument.

Original Calibration of H_2CO at Picarro

In 2018, Picarro determined the scaling of the G2307 instrument with a combination of approaches to anchor it most closely to trusted scales. The spectroscopy of Saha et al. (2007) provided guidance as to the frequency at which formaldehyde could be measured, and the relative line strength of the peaks, as equivalent features were not present in the HITRAN database at high enough resolution. The scaling determined by Saha et al. was checked in 2018 relative to a trusted Airgas-sourced tank owned by NASA

Goddard ("Airgas B")[2]. The measured value of 539.5 ppb was adjusted up by 1.0625x to account for matrix effects (see **Measurements in Air vs. Nitrogen Balance** later in this document) to give 573.2 ppb, which was about 3.5% low relative to the last measurement of the tank on NASA's UV instrument (593.4 ppb), but about 8% low relative to NASA's IR instrument (620 ppb). Because the Picarro agreed very closely with one of the two instruments, and because there was no clear answer as to which of the two NASA instruments was more accurate, Picarro retained its scaling factor. Picarro G2307 instruments produced since then are tied to the original spectroscopy with an expected scaling deviation well below 5% from unit to unit.

Scaling Differences, Round Robin Work

In 2019, the New York State Department of Environmental Conservation (NYSDEC) purchased two G2307 instruments as part of a process to evaluate their utility for real-time ambient air quality monitoring. Among their goals was to determine whether the scaling and accuracy of the G2307 agreed well with the scaling of the EPA TO-11A method for carbonyls—4,2-DNPH denuder cartridges (US EPA ORD, 1999). Initial testing of the two G2307 instruments by the NYSDEC revealed the need for a better understanding of why scaling differs relative to TO-11A, traceability of calibration for consistency between units, careful accounting of zero drift, and improvements in the water vapor calibration. Other documents will cover the latter two topics. This white paper focuses only on the first two efforts to improve the accuracy and traceability of the instrument's calibration scaling.

With initial indications from the NYSDEC intercomparison that the scaling of the Picarro G2307 might differ from that of the TO-11A method, Picarro worked with Apel-Riemer Environmental and NYSDEC to coordinate a Round Robin intercomparison with a number of experts in reactive species chemistry to characterize a known tank using a variety of different analytical techniques. In late 2019, Apel-Riemer sent around tank CC508579, owned by Eric Apel, NCAR, but filled originally by Apel-Riemer, to a series of labs for characterization.

---

[2] We note that the scaling adopted during development in 2018 as part of a comparison against a trusted Airgas tanks is highly consistent with observations on Picarro instruments using recently-produced Airgas formaldehyde blends. This suggests both that the Airgas blends are being produced consistently with time, and that Picarro instruments are highly consistent and accurate as they come off the line.

Table 1: Round Robin results for a 1690 ppb H_2CO in N_2 cylinder produced and certified by Apel-Riemer Environmental (CC508579), owned by Eric Apel, UCAR. A 1.0625% adjustment factor is applied to Picarro measurements to account for an air vs. N_2 matrix measurement bias, as discussed later in this white paper.

Individual	Institution	Instrument, Technique	Adj. factor	Date	Measured H_2CO [ppb]	Adj. H_2CO [ppb]	% of certified
Riemer, D	Apel-Riemer	GC FID	1	1/2020	1690	1690	1.000
Bent, J	Picarro	G2307 CRDS LBDS2003	1.0625	2/2020	1369	1455	0.861
		G2307 CRDS LBDS2030	1.0625	2/2020	1385	1472	0.871
Furdyna, P	NYSDEC	EPA TO-11A 4,2-DNPH	n/a	6/2020	1633	1633	0.966
Fried, A	CU INSTAAR	TDLAS	n/a	2020	1645	1645	0.973
Orlando, J; Tyndall, G	NCAR	FTIR	n/a	2020	1550	1550	0.917

Apel-Riemer certified the tank on a GC FID system at 1690 ppb ± 5% at their Florida facility in January 2020. Peter Furdyna at NYSDEC characterized the tanks with 4,2-DNPH cartridges, and used a mass flow controller positioned downstream of the sampler to confirm flow rate was controlled and at spec in April-June 2020. Offline HPLC measurements of the DNPH reported a value of 1633 ppb, about 3% below the certified value. Alan Fried at University of Colorado INSTAAR (Fried, McKeen, et al., 1997; Fried, Sewell, et al., 1997; Liao et al., 2021) measured the same tank to be 1645 ppb, also about 3% below the certified, and at the National Center for Atmospheric Research (NCAR) in Colorado, John Orlando and Geoff Tyndall (Bacher et al., 2001; Orlando et al., 2001) characterized a value of 1550 ppb, about 8% below the certified value. Picarro received the tank in February 2020 and ran the standard through a PFA tee to two instruments simultaneously, each of which retained its factory calibration curve. The two instruments reported values of 1369 and 1385 ppb. After correction for effects associated with the cylinder gas matrix (see **Measurements in Air vs. Nitrogen Balance**) the instruments report effective values of 1455 and 1472 ppb, about 14% and 13% below the certified value.

This initial intercomparison suggested that Picarro G2307 instruments read 8-15% low relative to instruments designed and run by multiple subject matter experts, and EPA methods. NYSDEC continued to perform intercomparisons with DNPH to confirm this relationship over a longer period at a deployed site under real-world conditions at the New York Botanic Gardens. Intercomparisons over 2020-2021 by the New York State Department of Environmental Conservation demonstrate the linear relationship between Picarro measurements and 4,2-DNPH cartridges. Measured with Picarro as the dependent variable, a slope of roughly 0.83 was determined on NYSDEC instrument LBDS2024, using its factory calibrated slope.

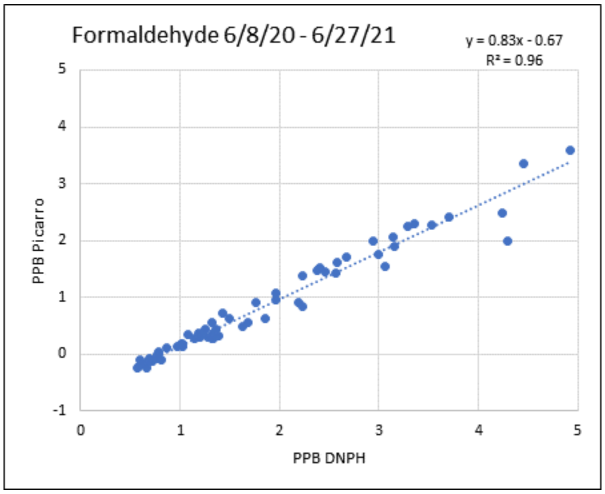


Figure 2: H_2CO comparison between 1-in-6-day collected 8-hour-integrated samples from period June 8, 2020 to June 27, 2021, using 4,2-DNPH cartridges (x axis) with averages from zero-corrected Picarro data (y axis). Zero air from an API 701H was provided to the instrument for minutes 0-7 of a given hour, and sample air from minutes 8-59. Stable values from minutes 4-7 were used to correct sample data taken over minutes 12-59. Hour-averaged G2307 data were then further averaged over an 8-hour window (y axis) for comparability with DNPH. N = 60 samples.

Furthermore, Picarro is aware of other anecdotal evidence that the G2307 instrument reads on order 10-30% low relative to other methods from users in North Carolina, California, Korea, and China. Though not every single intercomparison has found Picarro instruments to read low relative to other methods (e.g. Glowania et al., 2021), the great majority have. Accordingly, Picarro has decided to seek out gas standards that are widely accepted by the formaldehyde community to determine whether the instrument scaling requires adjustment. Furthermore, Picarro has chosen to anchor all future instrument builds to the newly determined scaling with a "Golden Analyzer" which will serve as a transfer standard, following the methods devised in previous efforts (Rella, 2017a, 2017b).

DETERMINING THE NEW SCALING OF THE G2307 GOLDEN ANALYZER

Comparison of the 2018 H_2CO G2307 Calibration against Reference Materials

For the calibration of the new instrument, we obtained three gravimetrically prepared standards of H_2CO in cleaned and treated N150-size Advanced Material Systems aluminum cylinders from Apel-Riemer Environmental Inc.. The tanks were prepared from paraformaldehyde stock from Alfa Aesar (97% / Stock No A11313 / L04737 / Lot No 10136752) and were characterized on a GC-FID by Apel-Riemer on September 10, 2021. They were shipped to Picarro shortly thereafter, where they were rapidly tested against the LBDS2047 Golden analyzer that had been chosen for the project. These standards were prepared with target concentrations of 0.5 ppm, 1 ppm, and 2 ppm H_2CO in N_2 with blend tolerances of ± 5%. Actual certified values by Apel-Riemer were 523.7, 1027, and 2032 ppb. For additional details on these tanks, see the **Appendix**.

Physical Setup

Tanks from Apel-Riemer were plumbed to the Picarro Golden directly using approximately 3m of PFA[3] tubing and a dedicated pre-flushed regulator per tank. The H_2CO standard in N_2 was allowed to run to the

---

[3] Picarro tested both PFA and Silconert-treated stainless steel for the direct delivery line on 3/5/2022, and found no statistically significant differences between the two materials.

instrument for as long as needed for the value to fully stabilize. While the value jumps to within 90% of its eventual value in a matter of seconds, the remaining asymptotic behavior takes up to four hours before a completely stable value can be read. Picarro refers to this long tail as a "pickling" period. We understand this to represent the time required for the tubing, regulator, and instrument parts to slowly adsorb microlayers of H_2CO on surfaces until the surfaces have adequately removed water and other polar molecules, and come into equilibrium with the partial pressure of the gas in the airstream.

In addition to the three tanks used for determining the span of the instrument, ambient air run through a DrieRite filter was used as the zero point for the calibration process. The instrument needs low levels of CH_4, H_2O, or H_2CO as spectroscopic reference points to ensure that the zero point does not drift during long periods of characterizing a clean zero, and while DrieRite removes H_2O and H_2CO, it preserves CH_4 in the matrix, preventing drift.[4]

Importantly, for the purposes of accuracy, Picarro does not blend down these three tanks to establish a calibration curve. We have observed that formaldehyde standards cannot be blended accurately, even in Silconerted regulators, so the three span values used in the assessment of the scaling/accuracy of the instrument are direct connections from each tank, as shown in the figure below (Fig. 3). The experiment was performed using a single instrument (S/N: LBDS2047), which we have designated as our "Golden" H_2CO Analyzer. We will describe how this Golden Analyzer is used in more detail in later sections.

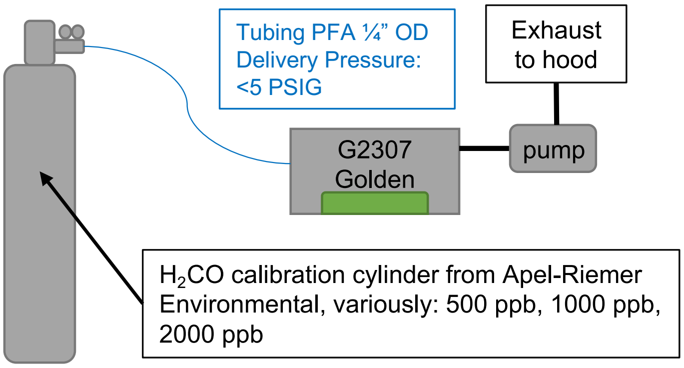


Figure 3: Schematic showing 'Golden' H_2CO analyzer creation setup. Each cylinder was measured one at a time with a direct connection to the inlet of the analyzer. The PFA tubing was 3m in length and the applied gas pressure was <5 psi.

---

[4] This form of drift relates to the ability of the spectrometer to correctly produce and read light at a known frequency. If one or more of the species defined in the model function at adequate concentrations is not present, the instrument is unable to lock onto a "frequency reference" value, and with time will begin to search for that frequency reference elsewhere in the spectrum. For the G2307, these values are: H_2O > 140 ppm **OR** CH_4 > 1.4 ppm **OR** H_2CO >300 ppb.

Two adjustments have been made in the intercomparison of the Golden instrument with the reference materials:

1) The instrument value at zero has been forced to zero. This reflects the strong understanding of the absolute value of the zero measurement on the instrument with a carefully zeroed gas stream.
2) Measured values for formaldehyde standards in N_2 have been adjusted upwards by a factor of 6.25%, i.e. for a given measured value of X, we report 1.0625X. This reflects the observed peak broadening effects on the LBDS analyzer, characterized in early 2018 as part of the original research and development effort on the instrument. We discuss this effect in the section below.

Measurements in Air vs. Nitrogen Balance

When relating absorption to concentration, it is the integrated absorption – the area under the entire absorption line – which is proportional to the concentration of the absorbing gas at a given temperature and pressure. This integral value is, however, significantly less precise than reporting the delta between the peak and an established baseline value, so Picarro analyzers generally take the latter approach for reporting concentrations. For an instrument intended to measure samples in an air background (~78% N_2; ~21% O_2), the scaling that relates absorbance (units of ppb/cm) assumes the presence of O_2 at those stated levels. When standards are provided to the instrument in N_2 instead, the additional ~21% mole fraction of N_2 has the effect of broadening the width of the formaldehyde peak more than the equivalent O_2 would, and lowering the peak value relative to the baseline, while retaining the same effective integral value. As a result, the G2307 Picarro reports slightly low when a standard in N_2 is sent to the instrument. This value can be adjusted upward to read correctly through multiplication by 1.0625, and accordingly could be said to read low by 5.88% (=(1-(1/1.0625)*100)). Picarro quantified this effect in early 2018 during the development of the LBDS platform.[5]

Accordingly, when running standards produced by Apel-Riemer to determine how different Picarro scaling is from Apel-Riemer scaling, we must run each standard for the requisite period and then adjust upward the mean stable value by 1.0625 before declaring this the adjusted “observed” value for the Picarro instrument. With this scaling bias accounted for, any remaining difference between the Picarro and the Apel-Riemer scaling can be attributed to historical differences in scaling factor choices.

Golden Scale Characterization Results

We plot the matrix-adjusted measurements of the Apel-Riemer tanks by the LBDS2047 Golden analyzer against the certified concentrations in Figure 4. The final calibration slope was calculated with a linear fit forced through the origin. The slope from the linear fit is **0.856,** indicating that the instrument reports *lower* than the prepared standard concentrations by a little more than 14%.

---

[5] As an example, consider a theoretical instrument which is perfectly calibrated relative to a well-prepared gas standard (reference material) of 500 ppb of formaldehyde in air. When the instrument is provided with this gas, a user would expect to observe 500 ppb exactly. Now consider a second “twin” standard produced with exactly 500 ppb H2CO in N2. When sent to the instrument under identical circumstances, the second tank would read low, at approximately 470.6 ppb. Multiplying this value by 1.0625 produces the correct value of 500 ppb.

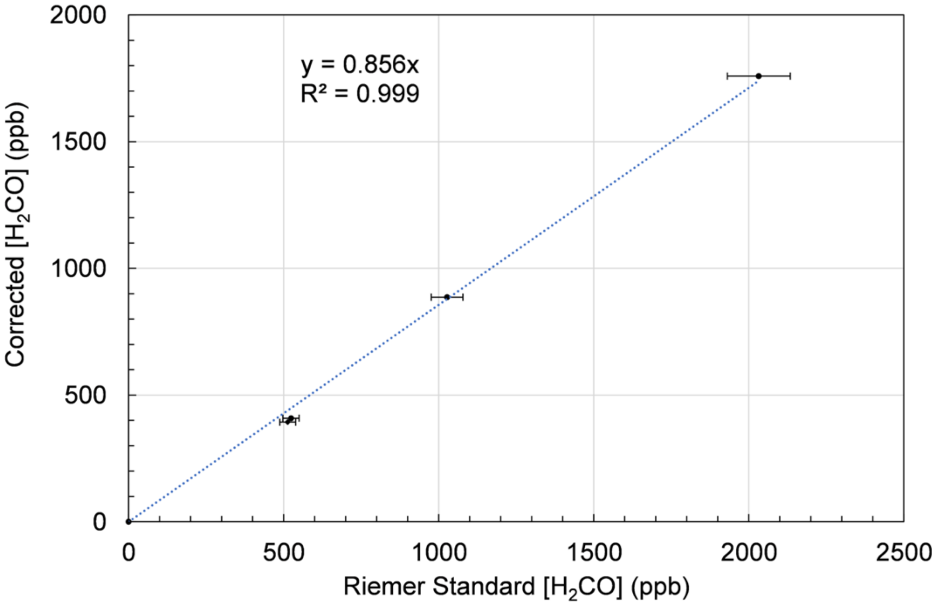


Figure 4: Measured cylinder values plotted against the assigned values for each cylinder (black symbols). The horizontal error bars indicate the analytical accuracy of ± 5% as determined by the manufacturer. Slope of 0.856 ppb/ppb and intercept of 0 ppb is the result of a linear fit (indicated by the blue dotted line) forced at the origin (by definition).

Table 2: The results of the comparison between Apel-Riemer reference materials and the LBDS2047 Golden analyzer.

Cylinder #	Manufacturer Certified Value [ppb]	Reported value on LBDS2047 [ppb]	Adjusted Reported Value (=1.0625x) [ppb]	StDev Raw [ppb]	Adj Rep / Certified
RR03056	523.7	384.8	408.8	0.972	0.781
RR03056 (2nd)	512.9	370.0	393.1	0.966	0.766
RR03068	1027	834.0	886.1	1.044	0.862
RR03454	2032	1655	1759	1.206	0.866
Drierite	n/a	0.345	0 (by definition)	0.959	n/a

Stability of Tanks Over Time

During the effort to characterize the Golden Instrument scaling, Picarro observed that Riemer tanks **RR03068** and **RR03454** displayed highly linear concentrations with respect to each other and zero, well within the 5% blend tolerance provided by the manufacturer relative to a slope fitted through zero. The third tank, **RR03056,** fell significantly below the value expected (~441 ppb) based on a fit to the two higher-concentration tank, at a N_2-matrix-adjusted value of 408.9 ppb. In November of 2021, we requested that tank **RR03056** be rerun by Apel-Riemer and sent the tank back for analysis. Upon reanalysis, Apel-Riemer found the value to have dropped by **10.8 ppb** to 512.9 ppb and recertified the tank at this new value. When the newly recertified tank was received at Picarro, the value had fallen yet further to a perceived adjusted value of 393.1 ppb on the Golden.

Gaseous standards of H_2CO are known to be challenging to store, so it is not surprising that we might observe a shipping-related loss that could be detected on both analytical techniques. We believe this effect is most likely tied to temperature more than time, as two flights over two weeks seem to have caused a drop of 15.7 ppb as observed on the Picarro instrument, adjusted relative to the Riemer scale. If a drop of **7.85 ppb** per shipment implies the original value of the tank was closer to the value stated by Apel-Riemer, with an estimated value of ~417 ppb, roughly at 0.8x, still about 7% low relative to the other standards. Because Picarro cannot avoid this effect and cannot defend this estimated extrapolation more than our observed numbers, we choose to retain both observed values for **RR03056**, despite the fact that these values fall outside of the 5% threshold certified by the manufacturer relative to an assumed relationship implied by a fit to the two higher-concentration tanks. We do this because we cannot reasonably suggest another approach that avoids temperature cycling or other travel-related effects, and because the biasing effect associated with this third tank is small relative to the other two tanks. We suspect anecdotally that lower concentration tanks are less stable than higher concentration tanks. This may imply a concentration-independent, possibly surface-area-dependent, degradation process that becomes more observable in lower concentration tanks. As of the writing of this document, Picarro does not have enough data to define the phenomenology of this effect. Since **RR03056** changed with time, we choose not to reject either value. Instead, we decided to factor both measurements into the fit.

Rescaling Summary

Accordingly, we find a slope of 0.856 to best relate the observed to the certified values. Thus, our new Golden Analyzer will be scaled upwards by 1.168 (=1/0.856) relative to the scaling factor determined in 2018. This factor is applied not to the original factory scaling, but to the InstrCal.ini file slope value for H_2CO for each G2307 analyzer for consistency with previous approaches. This 1.168 adjustment will also be applied to all future builds of the G2307 instruments for consistency with the Golden Analyzer.

Picarro adopts this new higher scaling factor because it aligns with:

1) A highly trusted calibration provider for formaldehyde and other reactive species in the US (Apel-Riemer Environmental), whose standards are used extensively by the US EPA and others (Gilpin et al., 1997; Liao et al., 2021).
2) The US EPA reference method for formaldehyde, which has been heavily vetted and described in multiple iterations of TO-11A guidance documentation (US EPA ORD, 1999).
3) Other instruments and analytical techniques from subject matter experts discussed above.

Table 3: Calibration adjustment of the Golden Analyzer.

Method	Calibration Adjustment Factor
Direct Cylinder Calibration Relative to Original Scaling	**1.168**

Conclusion: Picarro applies an upward adjustment of 1.168 to the original InstrCal scaling of the G2307 instruments on the LBDS2047 Golden Analyzer.

APPROACH FOR CALIBRATION OF INDIVIDUAL INSTRUMENTS

Overview

Because H_2CO is a fairly reactive gas, with atmospheric lifetimes between a few hours and few days (see e.g. Liao et al., 2021 and Alvarado et al., 2020, respectively), it is difficult to prepare and deliver a constant and known concentration of H_2CO, a fact which complicates individual instrument calibration. The desire to have all instruments measuring the same value further complicates the situation. We have therefore selected the following calibration approach:

1. Carefully calibrate a 'Golden' instrument. We have described this process in the preceding sections.
2. Use this 'Golden' instrument as a transfer standard to cross-calibrate each instrument built at the factory. Using the 'Golden' instrument as the calibration reference, rather than the gas preparation system, means that the stability and repeatability requirements for sample preparation can be relaxed dramatically.
3. Ensure the calibration of the 'Golden' instrument over time by checking the calibration of this instrument with a non-reactive proxy gas with an adjacent peak measured on the same laser. In the case of H_2CO, this proxy gas is CH_4, a greenhouse gas with atmospheric lifetime of well over 10 years, and calibration traceability through WMO and other primary gas scales.

DUT vs. Golden comparison methods

In Figure 5 we show the how a device under test (DUT)—i.e. an instrument being manufactured and tested for a customer—is compared against the Golden Analyzer used as a transfer standard. In the testing for this study, we used two mass flow controllers (MFCs), with full scales of 500 SCCM and 10 SLPM respectively. One MFC controlled the H_2CO flow from a custom blend of gas (here ~650 ppb) and one controlled a flow of ultra-high purity zero air diluent. The flows through the mass flow controllers (MFC) are configured to produce a set of concentration challenges to both instruments simultaneously. Because the two instruments shared a single gas preparation system, unavoidable losses of H_2CO in this system affected each instrument in an essentially identical way. Importantly, the two instruments must just agree well, not accurately reproduce an absolute value based on the blending performed by the mass flow controllers, since we cannot deliver a reliably accurate value of formaldehyde through a mass flow controller.

We installed an open split to allow us to produce a larger combined range of flows, keep inlet pressure reasonably constant, and to prevented mass flow controllers from actively fighting the flow rate of the analyzers (~430 SCCM each). We ensured the blending line between the join and open split tees was a minimum of 3' long to encourage mixing of the standard and diluent. A dividing tee downstream of the sample outlet of the open split provided near-identical flow conditions for the Golden and the DUT, and we cut tubing going to each instrument to the same length to minimize any changes relating to loss- or pressure-related effects.

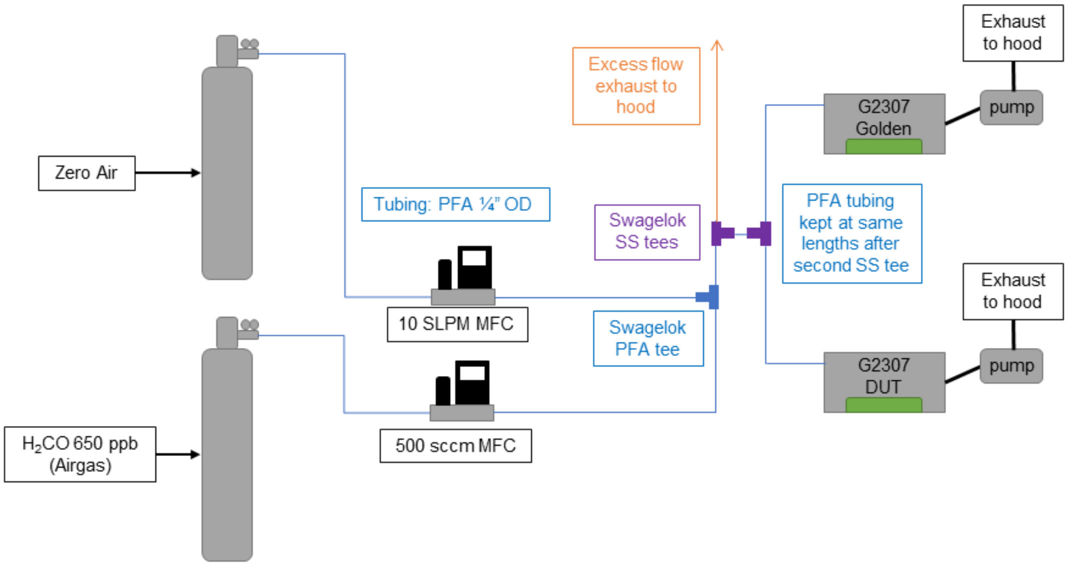


Figure 5. Cross-calibration setup used to transfer the calibration of the 'Golden' instrument onto the DUT.

Method Detection Limit Analysis

A partial typical time series showing the resulting concentration of the H_2CO is shown below (Figure 6). The testing program targets a series of 11 5-minute steps, here from roughly 135 to 8 ppb[6] and back, repeating the cycle three times to meet the requirements of the method detection limit procedure defined in Section 5.3 of SEMI (2014). The method detection limit in the aforementioned guide is defined as the *"level at which the errors in the measurement method become large enough such that the preset maximum acceptable risk [defined as 0.13%] of seeing the quantified level, when none of the contaminant in question is present in the sample, is exceeded."*

---

[6] These values are a function of the tank concentration, and may change very slightly as new testing tanks with slightly different certifications are used in the future.

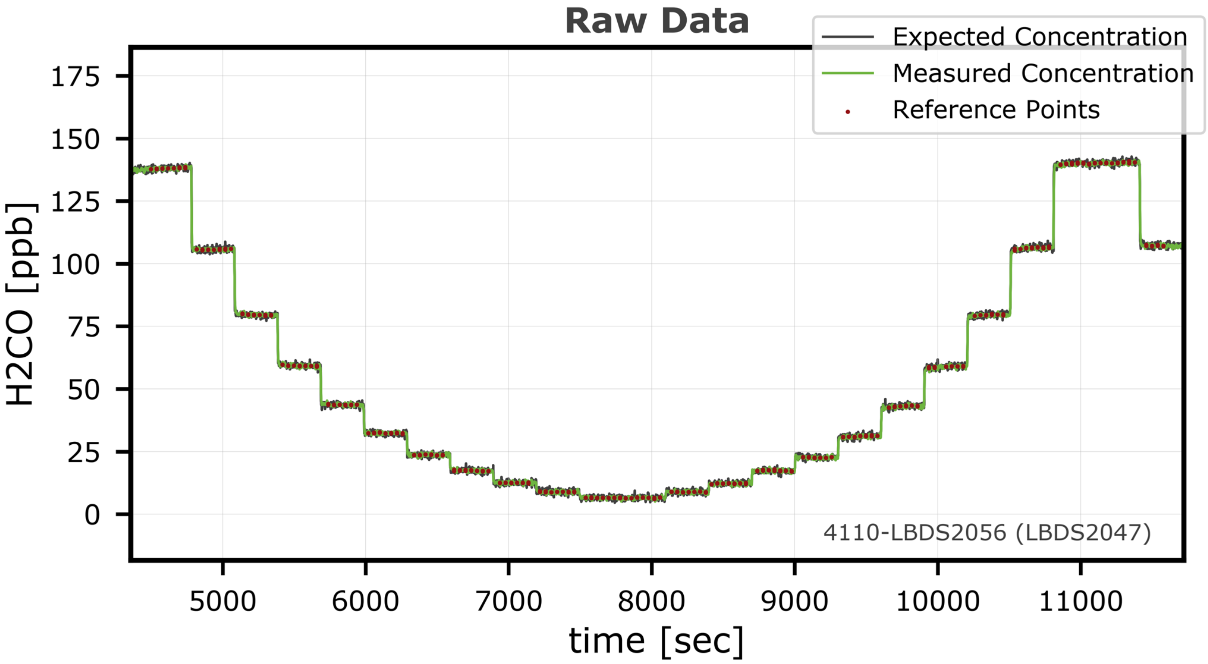


Figure 6: Time series for the Golden (LBDS2047 in black) and the DUT (LBDS2056, in green). The reference points (shown in red) are selected from both time series to meet minimum criteria for stability. These points are used in further analysis.

From that time series, the testing program identifies and flags a transition period during which the concentration is changing. It then averages the stable portion of each plateau for both the DUT and Golden analyzer. For each down-up cycle, the ascending and descending plateaus are averaged to ensure any memory effects do not bias the result downward or upward systematically.

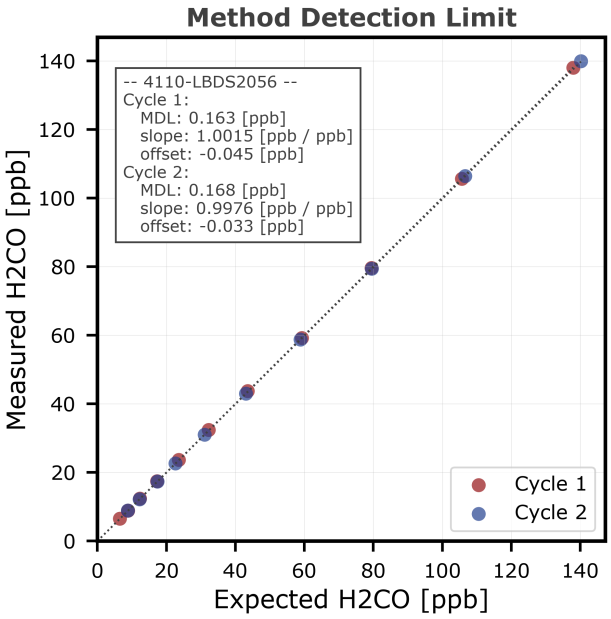


Figure 7. Calibration data derived from Figure 6. Differences from cycle to cycle the on value achieved relate to the repeatability of mass flow controller set points over short windows of time.

The average H_2CO value for each targeted concentration is then plotted for the DUT vs Golden instruments, shown above in Figure 7 as “Measured H2CO” and “Expected H2CO”, respectively. The method detection limit (MDL) is determined using an weighted least squares approach to assess the 3-sigma upper confidence limit at X=0 following the methods outlined in SEMI (2014).

Calibration Testing

From the same data used to determine the MDL, we can determine how the native calibration of the DUT instrument and the Golden differ by plotting the correlation of the two, as shown in Figure 8. In order for the instrument to pass this test, it must fall within 1.5% of the slope of the Golden instrument. The top left panel shows that the device under test, LBDS2056, comes well within that criterion at a relative slope of 1.0005, or 0.05% higher than the golden. Each residual from the golden must fit within the limits defined in the left bottom plot, less than the square root of the summed squares of 0.75 ppb and 1% of the reading.

Linearity Testing

If the instrument passes the calibration test, it must then pass a linearity test. The inverse of the slope of the DUT is applied to the DUT to align the two instruments to the same scaling, as seen in Figure 8, top right panel. The linearity of the instrument is then assessed using the methods outlined by SEMI (2014) and IEC (2010), with a Picarro-defined criterion that the agreement at each point should be better than 1% of the span range of the Golden instrument for the experiment (here 135 ppb), as seen in the bottom right panel.

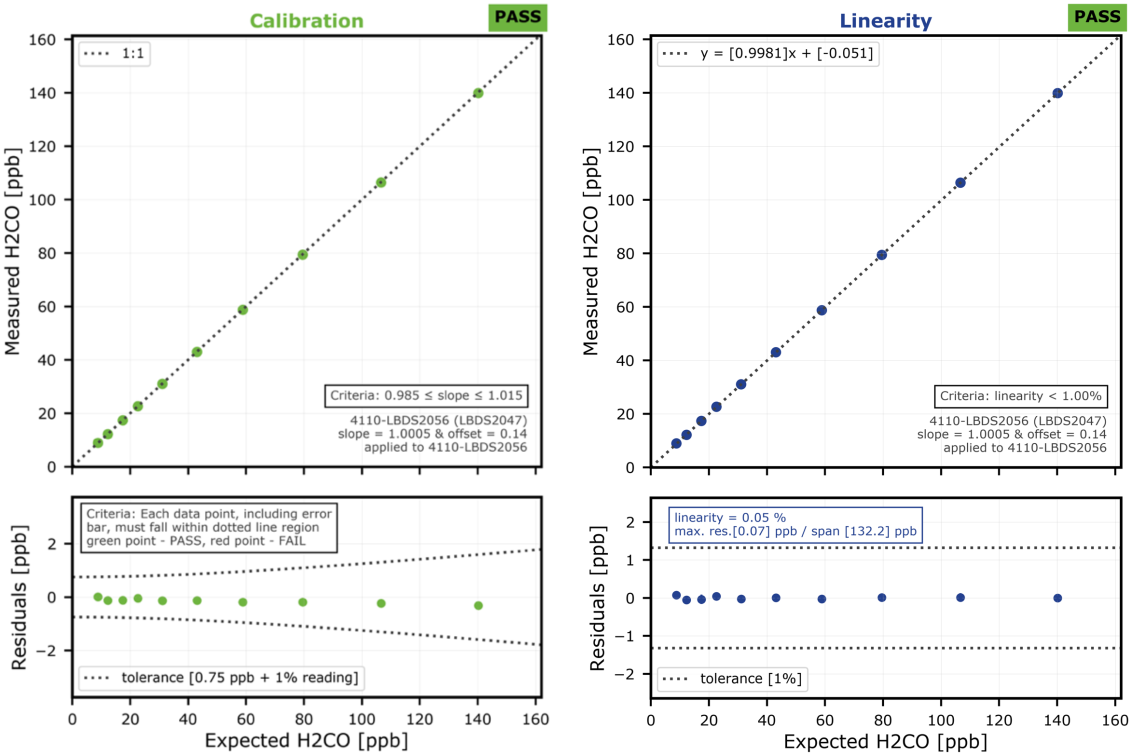


Figure 8: [Left, Top] Correlation plot of measured on the device under test (Y, LBDS2056) vs Golden (X, LBDS2047). Here, an absolute difference of 1.0005, or 0.05% was observed between the two slopes. **[Left, Bottom]** Residuals from the 1:1 line, each of which must be less than the square root of the summed squares of 0.75 ppb and 1% of the reading. [Right, Top] With the slopes adjusted to match each other for linearity testing ***only***, the deviation tolerance of 1% [Right, bottom] is assessed.

Stability of the Calibration Over Time

It is important to ensure that the calibration of each individual instrument will be stable between calibration verification events, which may occur as infrequently as once per year. It is very difficult to keep H_2CO standards stable for long periods of time. Permeation tubes and gravimetrically prepared cylinders of H_2CO generally have a certified life of 6-12 months, even though the balance gas is N_2. We therefore have little direct evidence of the stability of our H_2CO instruments over time using reference materials. There exists, however, ample evidence of the stability of the calibration of CRDS spectrometers over time periods of months and years. In Yver-Kwok et al. (2015), 47 Picarro CRDS instruments that measure CO_2, CH_4, and CO were studied, including 15 CRDS instruments that were calibrated in the field using highly accurate greenhouse gas standards. It was found that the calibration slope drifts typically about 0.1% / year; the largest drift observed overall was 0.3%. Because the CRDS spectrometers in these greenhouse gas instruments are functionally identical to the spectrometers used to quantify H_2CO, we may then conclude that the H_2CO calibration slope in the G2307 should exhibit a similar level of stability (<0.5%) over life. Because we expect the instrument stability to exceed our ability to deliver a known concentration of H_2CO, we recommend a yearly calibration validation, using either a trusted H_2CO standard or, more simply, the non-reactive CH_4 surrogate gas. There is no need to perform a true calibration in which the calibration slope is changed per the results of a direct H_2CO calibration experiment. For best absolute accuracy <10 ppb, consistent with the method used by NYSDEC, we do however recommend monitoring and correcting for small amounts of zero drift (on order <1 ppb) using scrubbed zero air. This process will be discussed in greater detail in subsequent documents.

Adjusting Existing Instruments to the New Scaling

Picarro advises users not to adjust the calibration of their instrument manually, but to seek assistance on this process from support@picarro.com. Picarro can perform the adjustment in a way that both ensures correct application, and which notes the scale adjustment in our records. Users may also return their analyzers for a recertification and rescaling at HQ for a fee if they are not comfortable having this change made by Support remotely.

Please contact support@picarro.com if you need assistance with this process, or to check whether your analyzer reflects the new, or original, scaling factor.

SUMMARY

In this white paper, we have demonstrated that the Picarro G2307 LBDS instrument has had a native scaling factor that causes it to read concentrations of formaldehyde about 10-20% low relative to other instruments and analytical techniques. Accordingly, we have used a set of trusted standards prepared by Apel-Riemer Environmental, which agree closely with the scaling of other techniques, to determine a new scaling relationship for the G2307 instrument platform. **We determine, based on a fit to three cylinders and a zero, that Picarro G2307 instrument scaling should be adjusted up by a factor of 1.168.** This adjustment is applied to a "Golden Analyzer" which will be responsible for maintaining this scaling. The stability of the Golden Analyzer will be verified periodically (about once per quarter), using both reactive H_2CO and a non-reactive surrogate gas, CH_4. All future G2307 instruments will receive this same 1.168 InstrCal scaling, and will be checked against the Golden for traceability. **This new adjustment factor has been applied to all G2307 analyzers from LBDS2060 onward**. Any instruments which disagree with the Golden by more than 1.5% in their accuracy, or which fail to meet the criterion of a 1% linearity check will be sent back to manufacturing to address performance issues, as instruments should be highly consistent given the fundamental principles of the build.

REFERENCES

Alvarado, L. M. A., Richter, A., Vrekoussis, M., Hilboll, A., Kalisz Hedegaard, A. B., Schneising, O., & Burrows, J. P. (2020). Unexpected long-range transport of glyoxal and formaldehyde observed from the Copernicus Sentinel-5 Precursor satellite during the 2018 Canadian wildfires. *Atmos. Chem. Phys.*, *20*(4), 2057–2072. https://doi.org/10.5194/acp-20-2057-2020

Bacher, C., Tyndall, G. S., & Orlando, J. J. (2001). The Atmospheric Chemistry of Glycolaldehyde. *Journal of Atmospheric Chemistry*, *39*(2), 171–189. https://doi.org/10.1023/A:1010689706869

Bent, J. (Picarro I. (2022). *Rapid Analyzer Validation Using a Traceable Surrogate Gas Approach,*.

Fried, A., McKeen, S., Sewell, S., Harder, J., Henry, B., Goldan, P., et al. (1997). Photochemistry of formaldehyde during the 1993 Tropospheric OH Photochemistry Experiment. *Journal of Geophysical Research: Atmospheres*, *102*(D5), 6283–6296.

Fried, A., Sewell, S., Henry, B., Wert, B. P., Gilpin, T., & Drummond, J. R. (1997). Tunable diode laser absorption spectrometer for ground-based measurements of formaldehyde. *Journal of Geophysical Research: Atmospheres*, *102*(D5), 6253–6266.

Gilpin, T., Apel, E., Fried, A., Wert, B., Calvert, J., Genfa, Z., et al. (1997). Intercomparison of six ambient

[CH2O] measurement techniques. *Journal of Geophysical Research: Atmospheres*, *102*(D17), 21161–21188. https://doi.org/https://doi.org/10.1029/97JD01314

Glowania, M., Rohrer, F., Dorn, H.-P., Hofzumahaus, A., Holland, F., Kiendler-Scharr, A., et al. (2021). Comparison of formaldehyde measurements by Hantzsch, CRDS and DOAS in the SAPHIR chamber. *Atmos. Meas. Tech. Discuss.* https://doi.org/10.5194/amt-2021-10

IEC. (2010). *IEC 61207-1:2010, Expression of performance of gas analyzers*.

Liao, J., Wolfe, G. M., Hannun, R. A., St. Clair, J. M., Hanisco, T. F., Gilman, J. B., et al. (2021). Formaldehyde evolution in US wildfire plumes during the Fire Influence on Regional to Global Environments and Air Quality experiment (FIREX-AQ). *Atmos. Chem. Phys.*, *21*(24), 18319–18331. https://doi.org/10.5194/acp-21-18319-2021

Orlando, J. J., Tyndall, G. S., & Ceazan, N. (2001). Rate Coefficients and Product Yields from Reaction of OH with 1-Penten-3-ol, (Z)-2-Penten-1-ol, and Allyl Alcohol (2-Propen-1-ol). *The Journal of Physical Chemistry A*, *105*(14), 3564–3569. https://doi.org/10.1021/jp0041712

Rella, C. (Picarro I. (2017a). *SI2104: Traceable calibration of Hydrogen Sulfide (H2S)*. Retrieved from https://semi.picarro.com/support/library/documents/traceable_calibration_of_hydrogen_sulfide_h2s

Rella, C. (Picarro I. (2017b). *SI2205 and SI2306: Traceable calibration of Hydrogen Fluoride (HF)*.

Saha, S., Barry, H., Hancock, G., Ritchie, G. A. D., & Western, C. M. (2007). Rotational analysis of the 2v5 band of formaldehyde. *Molecular Physics*, *105*(5–7), 797–805. https://doi.org/10.1080/00268970601126767

SEMI. (2014). *C10-1109 (Reapproved 1114): Guide for Determination of Method Detection Limits*.

US EPA ORD. (1999). *EPA/625/R-96/010b, Compendium Method TO-11A, Determination of Formaldehyde in Ambient Air Using Adsorbent Cartridge Followed by High Performance Liquid Chromatography (HPLC) [Active Sampling Methodology]*.

Yver Kwok, C., Laurent, O., Guemri, A., Philippon, C., Wastine, B., Rella, C. W., et al. (2015). Comprehensive laboratory and field testing of cavity ring-down spectroscopy analyzers measuring H 2 O, CO 2 , CH 4 and CO. *Atmospheric Measurement Techniques Discussions*, *8*(4), 4219–4272. https://doi.org/10.5194/amtd-8-4219-2015

APPENDIX: CERTIFICATES

Attached are the Certificates of Analysis for the three Apel-Riemer H2CO cylinders used for this study.

Cylinder RR03056 Page 1 of 2

APEL-RIEMER ENVIRONMENTAL, INC
REFERENCE GASES AND ATMOSPHERIC CHEMISTRY

Certificate of Analysis

Gas-phase Calibration Standard

This gas-phase standard is intended to be used as a reference material for the calibration of instruments.

Statement about preparation and traceability:

Standards are gravimetrically prepared in high-pressure aluminum cylinders (Luxfer, Inc., Riverside, California; Catalina Cylinders, Inc, Garden Grove, California; or Advanced Material Systems, Taoyuan, Taiwan). Cylinders are cleaned and treated to eliminate contamination and ensure inertness. Standards are prepared in N150 cylinders (~4000 Liters calibration gas), N033 cylinders (~800 Liters calibration gas), or N006 cylinders (~125 Liters calibration gas) at a pressure of 2000 psi of ultra-high purity nitrogen or ultra-zero air. Valves are high purity stainless steel or brass (Ceodeux-Rotarex, Lintgen, Luxembourg) with a CGA350 fitting for nitrogen mixtures or a CGA590 for air mixtures. Pure compounds as liquids, gases, and solids are obtained from a number of sources. All lot numbers are cataloged. The gravimetric preparation is performed using calibrated microbalances (Mettler-Toledo, Columbus, Ohio) and microsyringes (Hamilton, Reno, Nevada and SGE, Austin, Texas) for measuring the compounds and cylinder balances (Mettler-Toledo, Columbus, Ohio) for the balance gas. Balances are calibrated with NIST traceable weights.

We prepare each cylinder individually. Accuracy is better than +/- 5%. Analysis confirms the accuracy of the gravimetric preparation. We use a series of NIST, NIST-traceable, NPL, and in-house gravimetric standards to perform the instrument calibrations.

Stability varies depending on the compound, concentration, and cylinder size. Many compounds are stable for several years.

The calibration gas mixture in cylinder RR3056 is certified from the analysis date for 12 months.

Daniel D. Riemer, Ph.D.

1295 NW 163rd Street, Miami, Florida 33169 USA
Telephone: 786-925-6201 / Fax: 786-364-1591
Email: info@apelriemerenvironmental.com

Cylinder: RR03056
Cylinder Date: 2021/06
Valve: CGA350 SS 16/484241
Lot No.: 21226.3
Cylinder Pressure: 2000 psia
Analysis Date: September 10, 2021

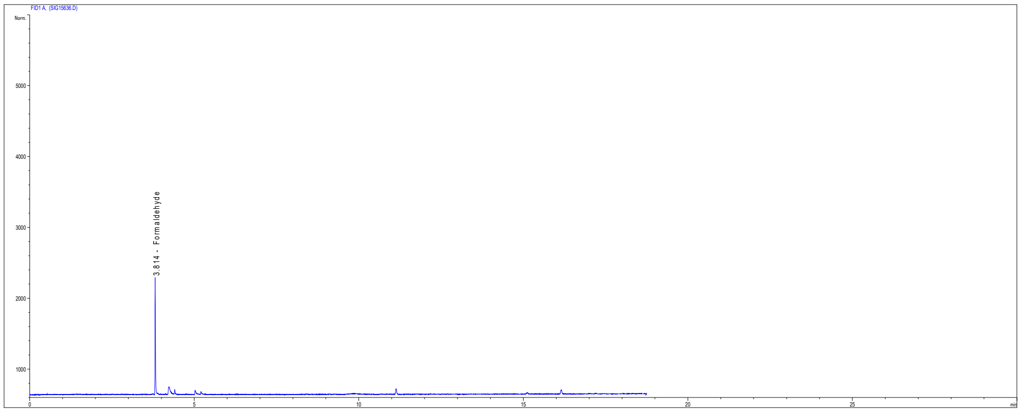
Single-component calibration mixture in nitrogen

Compound	CAS#	Concentration (ppb)	Uncertainty
Formaldehyde	50-00-0	523.7	±5%

Uncertainty is a conservative estimate of the combination of the uncertainties of the gravimetric preparation and analysis.

<u>Chromatogram</u>

100-meter DB-1, 0.25 mm id, 3.1 mL min^{-1} Helium carrier gas – constant flow
Temperature Program: 35°C, 3.5 min → 4.5°C min^{-1} → 180 °C, 6 min



APEL-RIEMER ENVIRONMENTAL, INC
REFERENCE GASES AND ATMOSPHERIC CHEMISTRY

Certificate of Analysis

Gas-phase Calibration Standard

This gas-phase standard is intended to be used as a reference material for the calibration of instruments.

Statement about preparation and traceability:

Standards are gravimetrically prepared in high-pressure aluminum cylinders (Luxfer, Inc., Riverside, California; Catalina Cylinders, Inc, Garden Grove, California; or Advanced Material Systems, Taoyuan, Taiwan). Cylinders are cleaned and treated to eliminate contamination and ensure inertness. Standards are prepared in N150 cylinders (~4000 Liters calibration gas), N033 cylinders (~800 Liters calibration gas), or N006 cylinders (~125 Liters calibration gas) at a pressure of 2000 psi of ultra-high purity nitrogen or ultra-zero air. Valves are high purity stainless steel or brass (Ceodeux-Rotarex, Lintgen, Luxembourg) with a CGA350 fitting for nitrogen mixtures or a CGA590 for air mixtures. Pure compounds as liquids, gases, and solids are obtained from a number of sources. All lot numbers are cataloged. The gravimetric preparation is performed using calibrated microbalances (Mettler-Toledo, Columbus, Ohio) and microsyringes (Hamilton, Reno, Nevada and SGE, Austin, Texas) for measuring the compounds and cylinder balances (Mettler-Toledo, Columbus, Ohio) for the balance gas. Balances are calibrated with NIST traceable weights.

We prepare each cylinder individually. Accuracy is better than +/- 5%. Analysis confirms the accuracy of the gravimetric preparation. We use a series of NIST, NIST-traceable, NPL, and in-house gravimetric standards to perform the instrument calibrations.

Stability varies depending on the compound, concentration, and cylinder size. Many compounds are stable for several years.

The calibration gas mixture in cylinder RR3068 is certified from the analysis date for 12 months.

Daniel D. Riemer, Ph.D.

1295 NW 163rd Street, Miami, Florida 33169 USA
Telephone: 786-925-6201 / Fax: 786-364-1591
Email: info@apelriemerenvironmental.com

Cylinder: RR03068
Cylinder Date: 2021/06
Valve: CGA350 SS 20/258942
Lot No.: 21226.2
Cylinder Pressure: 2000 psia
Analysis Date: September 10, 2021

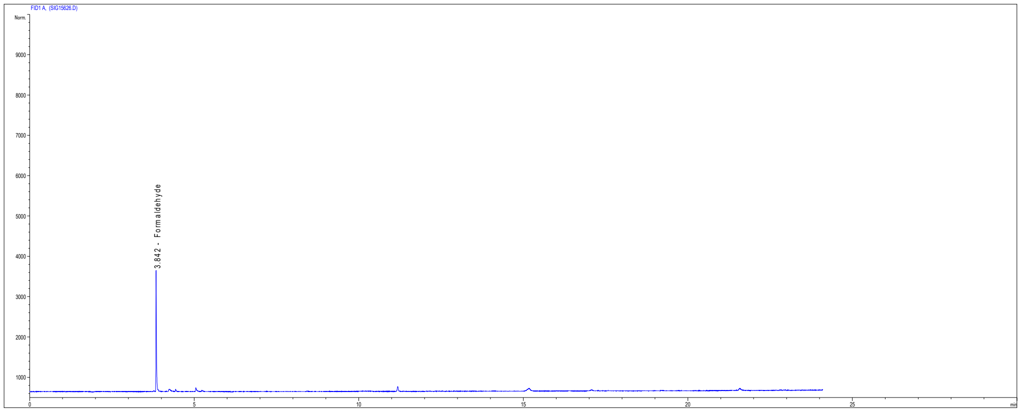
Single-component calibration mixture in nitrogen

Compound	CAS#	Concentration (ppb)	Uncertainty
Formaldehyde	50-00-0	1027	±5%

Uncertainty is a conservative estimate of the combination of the uncertainties of the gravimetric preparation and analysis.

Chromatogram

100-meter DB-1, 0.25 mm id, 3.1 mL min^{-1} Helium carrier gas – constant flow
Temperature Program: 35°C, 3.5 min → 4.5°C min^{-1} → 180 °C, 6 min



Cylinder RR03454 Page 1 of 2

APEL-RIEMER ENVIRONMENTAL, INC
REFERENCE GASES AND ATMOSPHERIC CHEMISTRY

Certificate of Analysis

Gas-phase Calibration Standard

This gas-phase standard is intended to be used as a reference material for the calibration of instruments.

Statement about preparation and traceability:

Standards are gravimetrically prepared in high-pressure aluminum cylinders (Luxfer, Inc., Riverside, California; Catalina Cylinders, Inc, Garden Grove, California; or Advanced Material Systems, Taoyuan, Taiwan). Cylinders are cleaned and treated to eliminate contamination and ensure inertness. Standards are prepared in N150 cylinders (~4000 Liters calibration gas), N033 cylinders (~800 Liters calibration gas), or N006 cylinders (~125 Liters calibration gas) at a pressure of 2000 psi of ultra-high purity nitrogen or ultra-zero air. Valves are high purity stainless steel or brass (Ceodeux-Rotarex, Lintgen, Luxembourg) with a CGA350 fitting for nitrogen mixtures or a CGA590 for air mixtures. Pure compounds as liquids, gases, and solids are obtained from a number of sources. All lot numbers are cataloged. The gravimetric preparation is performed using calibrated microbalances (Mettler-Toledo, Columbus, Ohio) and microsyringes (Hamilton, Reno, Nevada and SGE, Austin, Texas) for measuring the compounds and cylinder balances (Mettler-Toledo, Columbus, Ohio) for the balance gas. Balances are calibrated with NIST traceable weights.

We prepare each cylinder individually. Accuracy is better than +/- 5%. Analysis confirms the accuracy of the gravimetric preparation. We use a series of NIST, NIST-traceable, NPL, and in-house gravimetric standards to perform the instrument calibrations.

Stability varies depending on the compound, concentration, and cylinder size. Many compounds are stable for several years.

The calibration gas mixture in cylinder RR03454 is certified from the analysis date for 12 months.

Daniel D. Riemer, Ph.D.

1295 NW 163rd Street, Miami, Florida 33169 USA
Telephone: 786-925-6201 / Fax: 786-364-1591
Email: info@apelriemerenvironmental.com

Cylinder: RR03454
Cylinder Date: 2021/07
Valve: CGA350 SS 15/582579
Lot No.: 21226.1
Cylinder Pressure: 2000 psia
Analysis Date: September 10, 2021

Single-component calibration mixture in nitrogen

Compound	CAS#	Concentration (ppb)	Uncertainty
Formaldehyde	50-00-0	2032	±5%

Uncertainty is a conservative estimate of the combination of the uncertainties of the gravimetric preparation and analysis.

Chromatogram

100-meter DB-1, 0.25 mm id, 3.1 mL min^{-1} Helium carrier gas – constant flow
Temperature Program: 35°C, 3.5 min → 4.5°C min^{-1} → 180 °C, 6 min