Supplementary Material

Atmospheric H_2 observations from the NOAA Global Cooperative Air Sampling Network

Authors: Gabrielle Pétron[1,2], Andrew Crotwell[1,2], John Mund[1,2], Molly Crotwell[1,2], Thomas Mefford[1,2], Kirk Thoning[2], Bradley Hall[2], Duane Kitzis[1,2], Monica Madronich[1,2], Eric Moglia[1,2], Don Neff[1,2], Sonja Wolter[1,2], Armin Jordan[3], Paul Krummel[4], Ray Langenfelds[4], John Patterson[5]

1.Cooperative Institute for Research in Environmental Sciences, CU Boulder, USA
2. NOAA Global Chemical Laboratory, Boulder, USA
3. Max-Planck-Institute for Biogeochemistry (MPI-BGC), Jena, Germany
4. Commonwealth Scientific and Industrial Research Organisation - Environment, Aspendale, Australia
5. Department of Earth System Science, University of California, Irvine, USA

S1. Limitations of NOAA GML 1988-2009 H_2 measurements on RGAs

Novelli et al. [1999] describes the NOAA H_2 flask air measurement procedure for 1988-1997. A few aspects of the program for the period 1988-2009 are summarized here to explain limitations in the older NOAA H_2 dataset and the decision to not convert older measurements to the current WMO recommended calibration scale. These limitations can broadly be categorized as 1) issues related to the non-linear response of the analyzers used for flask analysis, 2) instability in the underlying internal scale maintained by GML, and 3) lack of adequate electronic records to provide full transparency. These all impact the quality and internal consistency of the early data and the ability to retroactively convert the early data to the current WMO recommended H_2 in air calibration scale.

Insufficient instrument response characterization

Prior to 2009, NOAA GML used gas chromatography followed by hot mercuric oxide reduction (GC-HgO) and the UV absorption detection of the resulting elemental mercury for both standard air and flask air analyses of H_2. GML used commercial Reduction Gas Analyzer GC modules with HgO bed reduction gas detector from Trace Analytical Inc. (Menlo Park, California) and Peak Laboratories, LLC (Menlo Park, California). The NOAA RGA analyzers measured both H_2 and CO in the same chromatogram. Table S3 (further below) gives a list of the RGA instruments and working standards in service prior to the adoption of the GC-HePDD measurement technique.

The first instrument used, R2 (RGA3 GC module with RGD2 detector), was found to have a linear response for CO and H_2 over the range of mole fractions in the background atmosphere [Novelli et al.,1991, 1992]. However, Novelli et al. [1992] cautioned that the instrument absolute response and linearity were HgO bed dependent and could change over time.

After 1990, all new HgO bed detectors had non-linear responses for both CO and H_2 [Novelli et al., 1998; Novelli et al., 2003]. CSIRO and MPI H_2 measurement teams have reported similar results [Francey et al., 2003; Jordan and Steinberg, 2011].

In 1991, GML started using a suite of standards covering a range of CO mole fractions to create calibration curves during dedicated instrument response calibration episodes approximately bi-weekly [Novelli et al., 1998]. This approach was not adopted for H_2, likely due to a lack of standards with stable

H_2. Instead, for H_2 measurements, GML used a 1-point calibration strategy where the CO reference air tank, which brackets each sample aliquot, was value assigned for H_2 and used as the single H_2 working standard for calibrating flask air sample measurements. This strategy ignored the non-linear response of the detectors.

The non-linearity of the RGA3 response was assumed to be negligible over the narrow range of H_2 observed in background air samples from remote network sites. However, the impact of the non-linear response also depended on the H_2 working standards being themselves close to ambient H_2 mole fractions. In actuality, recorded H_2 assignments for the older working standards used for flask analysis ranged from 470 ppb to 644 ppb. This would give rise to persistent non-linearity induced biases on time scales of 6-18 months (the typical lifetime of the working standards) in the H_2 measurement records. GML did not characterize the non-linearity of the H_2 response of the RGAs so cannot retroactively correct for this effect. The biases are expected to be significant for some time periods leading the authors to caution against using the NOAA early H_2 data records.

Instability in the NOAA H_2 X1996 calibration scale

NOAA H_2 mole fraction measurements from 1988 - 2009 are traceable to an internal calibration scale (NOAA H_2-X1996) maintained by GML. This scale was defined by five gravimetric standards made in 1995/1996 (CC73198, CC86013,CA01310, CC86208, CC86259), covering the range 485 - 600 ppb H_2.

The X1996 scale was propagated to the five working standards (tanks ID with * in Table S1) used between 1988 and 1995 for flask air sample analyses by measurement against the gravimetric standards in 1996 [Novelli et al., 1999]. However, these post deployment calibrations could not assess the stability of the working standards during usage prior to 1996 so any drift occurring in the working standards prior to 1996 would be unaccounted for leading to potential biases in the very earliest records.

After 1996, the NOAA H_2-X1996 scale was maintained by bootstrapping secondary standards forward in time. In this method, each secondary standard was used to directly calibrate its successor. This method assumed no drift was occurring in either the initial secondary standard, nor in any subsequent secondary standard. While care was taken to use cylinders for secondary standards that did not display initial high drift of H_2, we now know that H_2 stability in air standards contained in aluminum cylinders is rare and growth of H_2 over time is much more likely. The bootstrap method is likely to have introduced long-term instability in the scale.

This strategy ignored the non-linear response of the detectors. The non-linearity of the RGA3 response was assumed to be negligible over the narrow range of H_2 observed in background air samples from remote network sites. However, this also depended on the H_2 working standards being themselves close to ambient H_2 mole fractions. In actuality, the working standards used for flask analysis often varied significantly from ambient background H_2 values. This would give rise to persistent non-linearity induced biases on time scales of 6-18 months (the typical lifetime of the working standards) in the H_2 records. GML did not characterize the non-linearity of the H_2 response of the RGAs so cannot retroactively correct for this effect. The biases are expected to be significant for some time periods leading the authors to caution against using the early H_2 data records.

Incomplete record keeping early on

There is no electronic record of any calibration and no recorded assigned value for R7 working standard AAL-17259. All R5 and R6 working standards have assignments on X1996 recorded back in June 2014, covering a wide range: 470-650 ppb. Only the later R5 standards (CC105928, CC71649) and R6 standards (CA06591, CC305198) have assignments with a linear drift coefficient. The other standards were assumed stable.

In addition to the other known limitations in the early implementation of the H_2 measurements, the lack of record keeping during the early years plays a role in the decision to not retroactively convert the early data to the current WMO recommended calibration scales. Documentation of decisions on standard value assignments, electronic records of raw data files for the instrument responses, and details of calibration hierarchy from the early records are often missing or lack sufficient detail. Unfortunately, this makes it impossible to recover the data, even within the larger uncertainties associated with the measurement issues discussed.

Examples of observed biases in the older NOAA H_2 measurements

Close in time analysis of CC119811on P2 in 2007 and 2008 against one of three SX standards (SX3540, SX-3523 or SX-3554) show a > 20 ppb spread in the derived H_2 (SI Figure 12), suggesting a strong non-linear response. The response of the P2 instrument was never fully characterized. However, Novelli et al. [2009] show results for eight tanks analyzed on P2 using one point or two point calibration compared to their results on H9. The one point calibration results show the larger biases, especially for tanks with H_2 furthest from the H_2 in the reference/standard (525 ppb): underestimation for tanks with H_2 below 525 ppb reaching close to -20ppb at 420 ppb and overestimation for tanks with H_2 above 525 ppb reaching +12 ppb at 593 ppb.

The responses of the R5 and R6 instruments were never fully characterized. However NOAA started the regular analysis of target air tanks on the MAGICC1 and MAGICC-2 systems in 2004. Results for target air tanks CC71583 (D) and CC1824 (H) are plotted in SI Figure 13 using different symbols and colors for different working standards. GC-HePDD measurements after 2008 show H_2 growing in both tanks. The earlier results on R5, R6 and P2 are scattered and suggest inconsistent assignments between the working standards, also likely including incorrect drift estimates. It is not robust to extrapolate a tank H_2 assignment based on available measurements on H9 a few or several years back in time as it is well known that the stability or growth of H_2 in high pressure aluminum cylinders can change over time.

S2. Same air comparison with CSIRO for NOAA historical H_2 data

In 1980, CSIRO GASLAB started GC measurements of CO_2, CH_4 and CO in air samples collected regularly at the Cape Grim Observatory. CSIRO switched to an RGA3-1 instrument from Trace Analytical in 1991 to measure CO and then also H_2. In 1992, CSIRO also started monitoring the RGA3-1 instrument response with a suite of 15 cylinders with (mostly stable) CO mole fractions spanning 20 - 400 ppb. To address the challenge of drifting CO and H_2 in most high pressure cylinders, in 1993, the CSIRO GASLAB started using "dilution experiments" of above ambient mole fraction tank air with known CO (and H_2) to CH_4 ratios with ultra pure zero air and tied the diluted air mixtures CH_4 assignments to a gravimetrically defined CH_4 calibration scale. They used the dilution experiments to periodically characterize the non-linearity of their GC-HgO instrument for CO and H_2.They found the instrument response was “significantly non linear” and of similar shape for both gasses (of the form $y=ax_2+bx+cx^d$, where x = peak height and a,b,c,d are estimated parameters from the response function fit) but for a while used a single response function for H_2 as they had too few stable H_2 standards outside of the ambient range [Francey et al., 2003].

The intercomparison of measurements by NOAA GML and CSIRO same air from the Cape Grim Observatory (1992-1998) showed significant (>2%) and trending biases [Masarie et al., 2001, Francey et al., 2003]. The non-linear response of the H_2 analytical system detector, the instability of H_2 standards stored in aluminum cylinders (commonly used for CO_2 and CH_4 standards) and the different calibration scales were presented as likely explanations for the observed time dependent biases.

S3. WMO/MPI-BGC X2009 H_2 calibration scale

To support advances in the understanding of the H_2 global budget, high quality and comparable observations are a non-negotiable requirement and should be anchored by a common stable calibration scale [WMO, 2007]. The Max Planck Institute (MPI) in Jena secured funding to support their laboratory work to investigate the stability of the H_2 mole fraction for reference air in various types of high pressure cylinders and to develop an accurate H_2 calibration scale. Jordan and Steinberg [2011] analyzed 100 air standards multiple times over a one to six year period on their GC-HgO instrument calibrated using multiple H_2 in real air standard gasses to fully describe the detector nonlinear response. They concluded that the H_2 mole fraction for reference air in steel and stainless steel cylinders did not drift significantly (< 1.5 ppb/yr). For aluminum cylinders however, they found a wide range of H_2 mole fraction drift rates (< 1.5 ppb/yr to > 20 ppb/yr) and drift behaviors (short term, ie. drift over a few months, to continued growth in H_2). The MPI X2009 scale became the official WMO scale for H_2 in 2011 [Jordan and Steinberg, 2011]. It is defined by thirteen standards (of which 12 are in stainless steel cylinders) with H_2 dry air mole fractions ranging from 139 ppb to 1226 ppb.

Once a CCL was established for H_2, experts from the WMO GAW recommended measurement laboratories adopt the WMO/MPI 2009 scale and develop procedures to track drifts in their standards and to appropriately characterize their instrument responses [WMO/GAW, 2014].

In 2007-2009, GML prepared 6 H_2 gravimetric standards ranging from 230 to 790 ppb in electropolished stainless steel cylinders (Essex Cryogenics, with tank IDs SX-#). Early results in GAW laboratories suggested H_2 was likely more stable in these cylinders than in aluminum cylinders. However, the new gravimetric mixtures differed by about +20 ppb compared to two H_2 secondary standards in aluminum cylinders GML used for the calibration of tertiary standards on the X1996 scale (Novelli, personal communication). In following years, GML continued using the 1996 gravimetric primary standards to define its internal H_2 calibration scale and also regularly measured the H_2 secondary standards against the stainless steel standards.

S4. MAGICC-3 reference air CA04145

To evaluate the stability of the reference air H_2 and the validity of the H_2 instrument response curve fit coefficients between MAGICC-3 instrument response calibration dates, we derive an H_2 assignment for the reference air cylinder for each instrument response calibration date (ratio of peak heights =1). For each MAGICC-3 reference air cylinder, we calculate its mean H_2 for the time period for which it was in use. The mean H_2 values for the 6 reference air cylinders used so far range from 542 and 583 ppb.

In Figure 3 we plot the deviation of each reference air cylinder assignment from its mean value as a function of the MAGICC-3 calibration date. The very first reference CA04145 air cylinder had the largest growth in its H_2 mole fraction: + 7.5 ppb in 5 months (~ 18 ppb/yr). The incremental increase between calibration dates is larger when the calibration becomes less frequent in late 2019. We apply a correction of 18 *(Δt) to flask analysis results on H8 between 11/6/2019 and 1/16/2020 with Δt being the difference between the flask analysis decimal date and the preceding response calibration decimal date (corresponding to calendar dates 11/6/2019, 12/4/2019 or 1/7/2020). For the period 3/26 to 8/1 2020 with the second reference air cylinder, H8 was more noisy and the increments in the reference air H_2 between response calibration dates jumped from -1 ppb to 1 ppb twice.

S5. References

Jordan, A., Damak F., ICOS Central Analytical Laboratories Flask and Calibration Laboratory, Quality Control Report 2022, 85p, available at: https://www.icos-cal.eu/fcl/qc-report [Last Accessed December 13, 2023].

Francey, R. J.; Steele, L. P.; Spencer, D. A.; Langenfelds, R. L.; Law, R. M.; Krummel, P. B.; Fraser, P. J.; Etheridge, D. M.; Derek, N.; Coram, S. A.; Cooper, L. N.; Allison, C. E.; Porter, L.; Baly, S. The CSIRO (Australia) measurement of greenhouse gases in the global atmosphere. In: Tindale, N. W.; Derek, N.; Fraser, P. J. editors, editor/s. Baseline Atmospheric Program Australia. Melbourne: Bureau of Meteorology and CSIRO Atmospheric Research; 2003. 42-53.
http://hdl.handle.net/102.100.100/191835?index=1

Masarie, K. A., et al. (2001), NOAA/CSIRO Flask Air Intercomparison Experiment: A strategy for directly assessing consistency among atmospheric measurements made by independent laboratories, J. Geophys. Res., 106(D17), 20445–20464, doi:10.1029/2000JD000023.

Novelli, P. C., Elkins, J.W., Steele, L. P., The Development and Evaluation of a Gravimetric Reference Scale For Measurements of Atmospheric Carbon Monoxide, J. Geophys. Res., 96, 13,109-13,121, 1991.

Novelli, P. C., Steele, L. P., and Tans, P. P., Mixing ratios of carbon monoxide in the troposphere, J. Geophys. Res., 97, 20,731-20,750, 1992.

Novelli, P. C., Lang, P. M., Masarie, K. A., Hurst, D. F., Myers, R., and W., E. J.: Molecular hydrogen in the troposphere: Global distribution and budget, J. Geophys. Res., 104, 30427–30444, 1999.

Novelli, P. C., Crotwell, A. M., and Hall, B. D., Application of Gas Chromatography with a Pulsed Discharged Helium Ionization Detector for Measurements of Molecular Hydrogen, Env. Sci. Technol. (43), 2431-2436, doi: 10.1021/es803180g, 2009.

WMO/GAW, 17th WMO/IAEA Meeting on Carbon Dioxide, Other Greenhouse Gases and Related Tracers Measurement Techniques (GGMT-2013) (Beijing, China, 10 - 13 June 2013) Edited by Pieter Tans and Christoph Zellweger, report 213, 2014. Accessible at https://www.uncclearn.org/wp-content/uploads/library/gaw_213_en.pdf. Last accessed November 10, 2023.

SI Tables

SI Table S1: H9 target tanks and the polynomial best fits to their calibration histories

Tank ID (fill)	Calibration date range on H9	t0	Assignment at t0 (ppb)	C1 (ppb/yr)	C2 (ppb/yr^2)	N	Residual standard deviation (ppb)	Fill date (location if known) (R=Refilled)
CC311842 (A)	2019-2022	2020.9878	478.6	0	0	8	0.32	2009-09-04 (NWR)
ND33960 (C)	2018-2022	2019.9289	529.5	0	0	11	0.43	2014-03-05 (NWR)
CC121971 (G)	2019-2022	2021.0834	546.5	0	0	9	0.30	2012-05-10 (NWR)
CA06194 (B)	2019-2022	2020.7726	578.4	0	0	10	0.49	2008-09-25 (NWR)
ND16439 (A)	2008-2015	2009.66673	635.9	0	0	9	0.54	2002-01-01 (R)
CA08247 (A)	2020-2022	2021.2483	675.1	0	0	7	0.73	2008-10-01 (NWR)
CA05278 (A)	2008-2014	2011.8239	675.2	0	0	7	0.56	2007-03-01 (MPI) (R)
CA05300 (A)	2008-2014	2011.8667	596.8	0.84	0	7	0.31	2007-03-01 (MPI) (R)
CC71607 (A)	2008-2021	2016.889	537.9	0.44	0	18	0.34	1991-10-01
CC73110 (A)	2008-2021	2016.1309	563.8	0.79	0	19	0.41	1990-01-01 (NWR, SM Luxfer)
CA04551 (F)	2012-2016	2014.9953	523.18	4.55	0	42	0.32	2011-12-21 (NWR)
CA07328 (A)	2008-2010	2009.2785	598.7	2.83	0	6	0.20	2006-10-02 (SM, grav blend)
CB10910 (B)	2018-2022	2019.8396	577.28	3.51	0	11	0.40	2016-02-18
CC71579 (F)	2008-2012	2011.3385	605.6	7.74	0	26	0.36	2008-09-19 (NWR) (R)
CA08145 (C)	2016–2017	2016.7627	646.5	27.2	0	20	0.48	2015-08-14 (NWR)
ALM-065166 (A)	2008-2022	2014.6308	659.0	0.26	0	8	0.69	2006-01-01
CC309852 (A)	2009-2019	2015.1105	227.5	2.23	-0.39	9	0.93	2009-10-01 (SM, grav blend)
CC309852 (A)*	20011-2019	2015.7837	226.8	1.66	-0.16	8	0.36	2009-10-01 (SM, grav blend)
CC327035 (C)	2019-2022	2020.7333	370.5	5.76	-0.48	10	0.23	2017-10-13 (NWR)

CA07339 (B)	2018-2022	2019.9513	365.0	4.777	-0.32	11	0.37	2010-03-01 (BLD, CO grav blend)
CA06827 (I)	2019-2022	2021.1466	433.5	1.91	-0.30	15	0.27	2018-11-09 (NWR)
CA06327 (D)	2019-2022	2021.3555	437.0	2.94	-0.56	16	0.22	2018-11-09 (NWR)
ND15749 (A)	2008-2022	2014.5413	563.6	0.40	-0.02	22	0.27	2001-01-01
CC310014 (B)	2018-2022	2019.6369	572.9	-0.03	0.19	26	0.24	2010-04-29 (NWR)
ND16443 (A)	2008-2022	2015.0192	604.6	0.45	-0.03	20	0.32	2001-01-01
ND17445 (A)	2008-2022	2014.9725	632.9	0.99	-0.07	22	0.46	2001-01-01
ND17435 (A)	2008-2022	2015.3295	686.9	0.47	-0.05	19	0.76	2001-01-01
CA05554 (B)	2010-2016	2014.7948	699.67	0.85	0.46	53	0.83	2009-10-23 (NWR)

* Alternative assignment when the tank first calibration result, 5 weeks after its fill date in 2009, is dropped from the fit.

SI Table S2: MAGICC systems target tanks and the polynomial best fits to their calibration histories

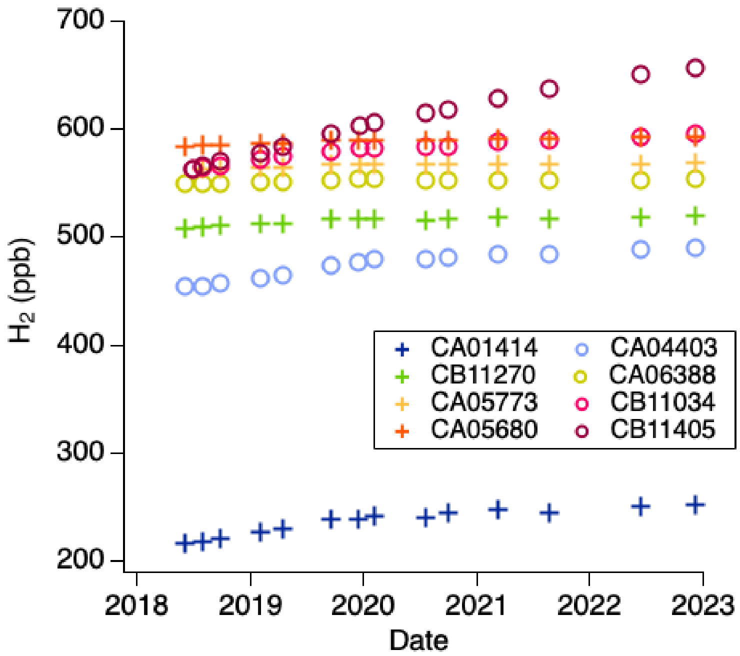
Tank ID (fill)	Calibration date range on H9	t0	Assignment at t0 (ppb)	C1 (ppb/yr)	C2 (ppb/yr^2)	N	Residual standard deviation (ppb)	Fill date (location if known)
CC1824 (H)	2009-2011	2010.1738	574.5	6.22	0	4	0.51	2006-07-06 (NWR)
CB08834 (B)	2011-2018	2015.6272	537.8	4.06	-0.50	10	0.57	2011-10-20 (NWR)
CC303036 (A)	2010-2017	2013.1491	588.3	21.31	0.47	10	0.44	2008-12-04 (NWR)
CB11143 (C)	2019-2022	2020.6759	534.7	1.91	0	9	0.54	2018-11-01 (NWR)
ALMX067998 (C)	2016-2022	2019.4574	542.1	0.62	0	13	0.28	2016-02-12 (NWR)
CB10292 (B)	2020-2022	2021.4553	597.4	0.95	0	5	0.44	2019-10-17 (NWR)
SX-1009237 (A)	2022-2023	2021.1697	526.5	0	0	2	0.24	2022-11-16 (BLD)

SI Table S3: List of instruments and reference air tanks used for H_2 in air sample measurements in GML

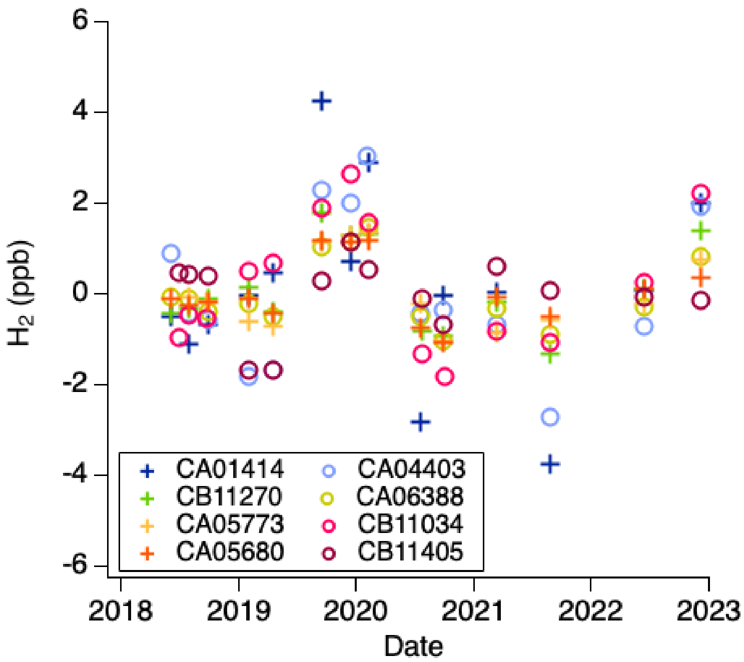
Tank Air Analysis						
Dates of operation	System	Instr ID	Model	Response	Secondary standard tank ID	Notes
1993-1997	rgd2	R2	RGD2	Linear	CC73110*, CC71607	No electronic records of calibration prior to 2001. Later used as TGT for H9.
1997-2006	rgd2	R7	RGA3	Non-linear	CC73110*, CC71607	
2006-2008	cocal-1	P2	PP1	Non-linear	CC119811	See SI Figure 12
Flask Air Analysis						
Dates	System	Instr. ID			Working standard tank ID	Notes
1988-1990	rgd2	R2	RGA3	Linear	AAL-17262, CC68734*	* H_2 was assigned against 1996 gravimetric standards and early data was reprocessed [Novelli et al., 1998].
1990-1995	carle	R4	RGA3	Non-linear	AAL-17269*, AAL-17270*, CC105871*	
1995-1997	carle	R7	RGA3	Non-linear	CC105871, AAL-17259	Assignments for later working standards were mostly inferred from earlier tanks, assuming no drift.
1997-2010	MAGIC C-1	R5	RGA3	Non-linear	CA02439,CA01493, CA02952,CA01777, CC61344, CA06593, CC105928, CC71649	
2004-2009	MAGIC C-2	R6	RGA3	Non-linear	CA02439, CA06527, CC68676, CA06591, CC305198	

SI Figures

SI Figure 1. H_2 calibration histories of eight MAGICC-3 working standards

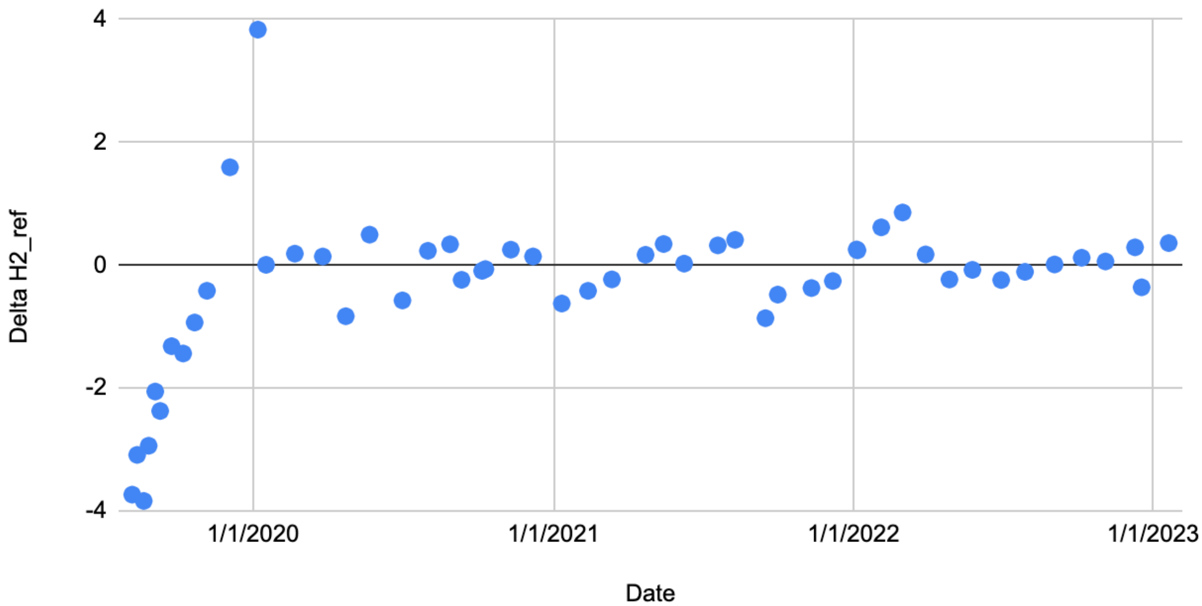


SI Figure 2: H_2 residuals from the calibration history trend function for eight MAGICC-3 working standards (see Table 3 of main paper)

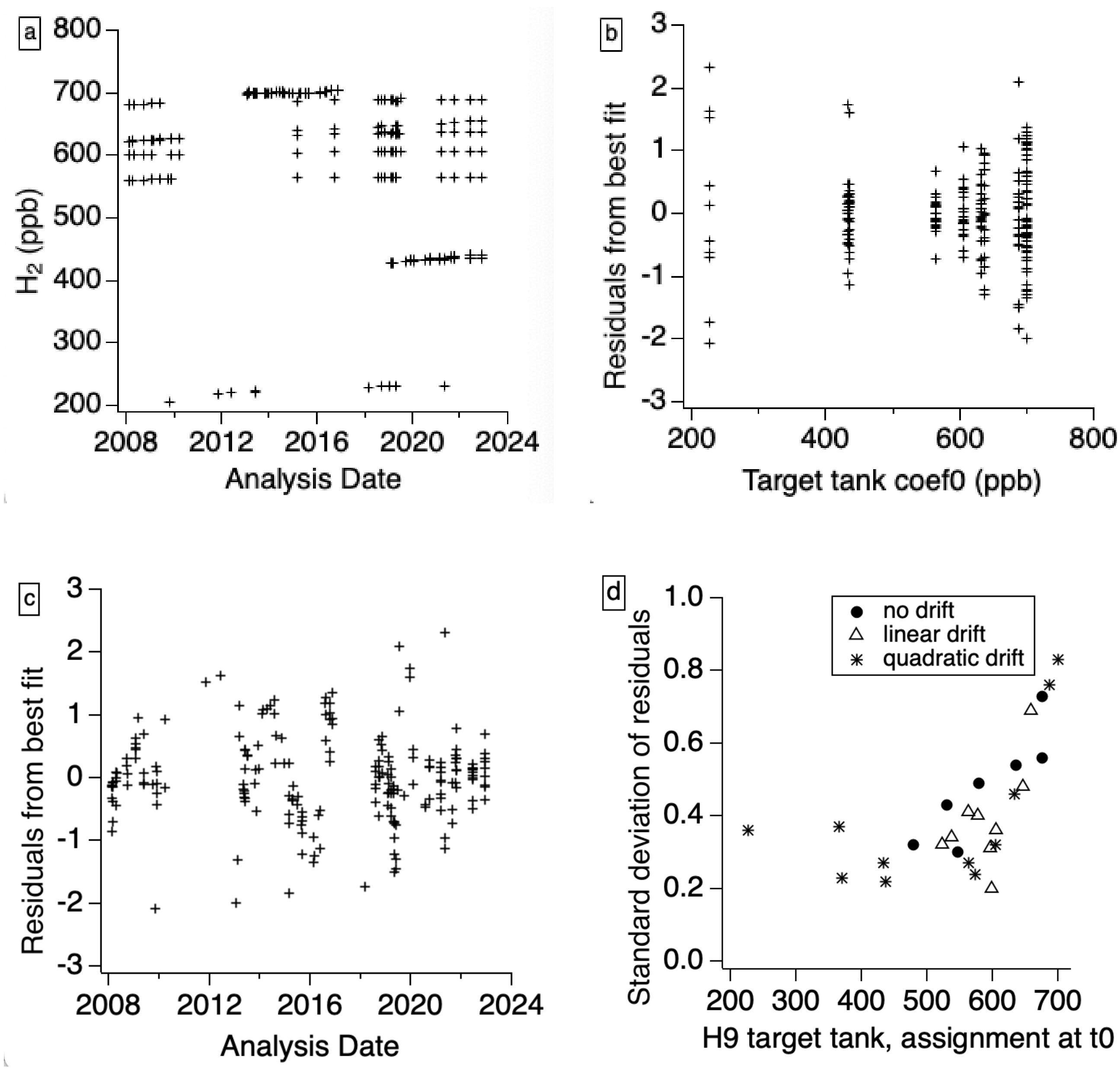


.

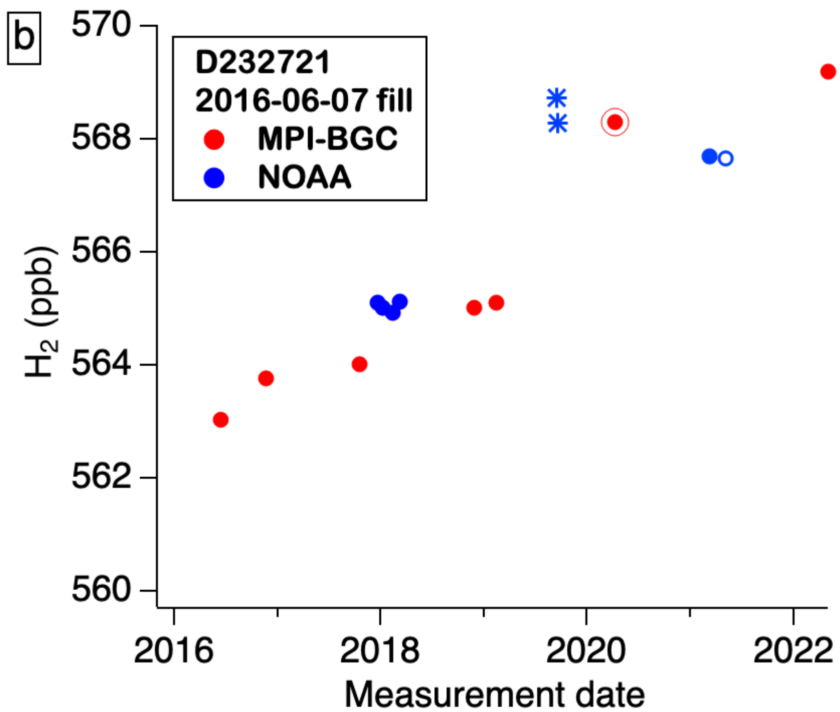
SI Figure 3. MAGICC-3 reference air deviation over time from mean. H_2 derived from response curves with x=1.

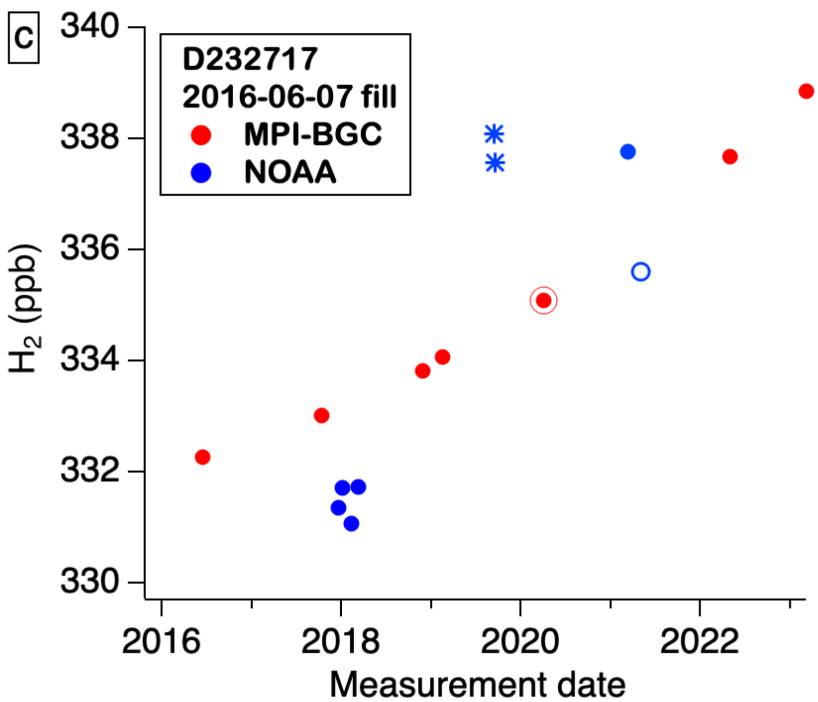


SI Figure 4. H9 Target tanks with quadratic polynomial fits to their calibration histories shown in plot a). Residuals from each tank best fit are shown in b) as a function of the initial assignment and c) as a function of the tank analysis date. d) Residuals standard deviation versus initial assignments (coef0) for all H9 Target tanks. All values are in ppb.

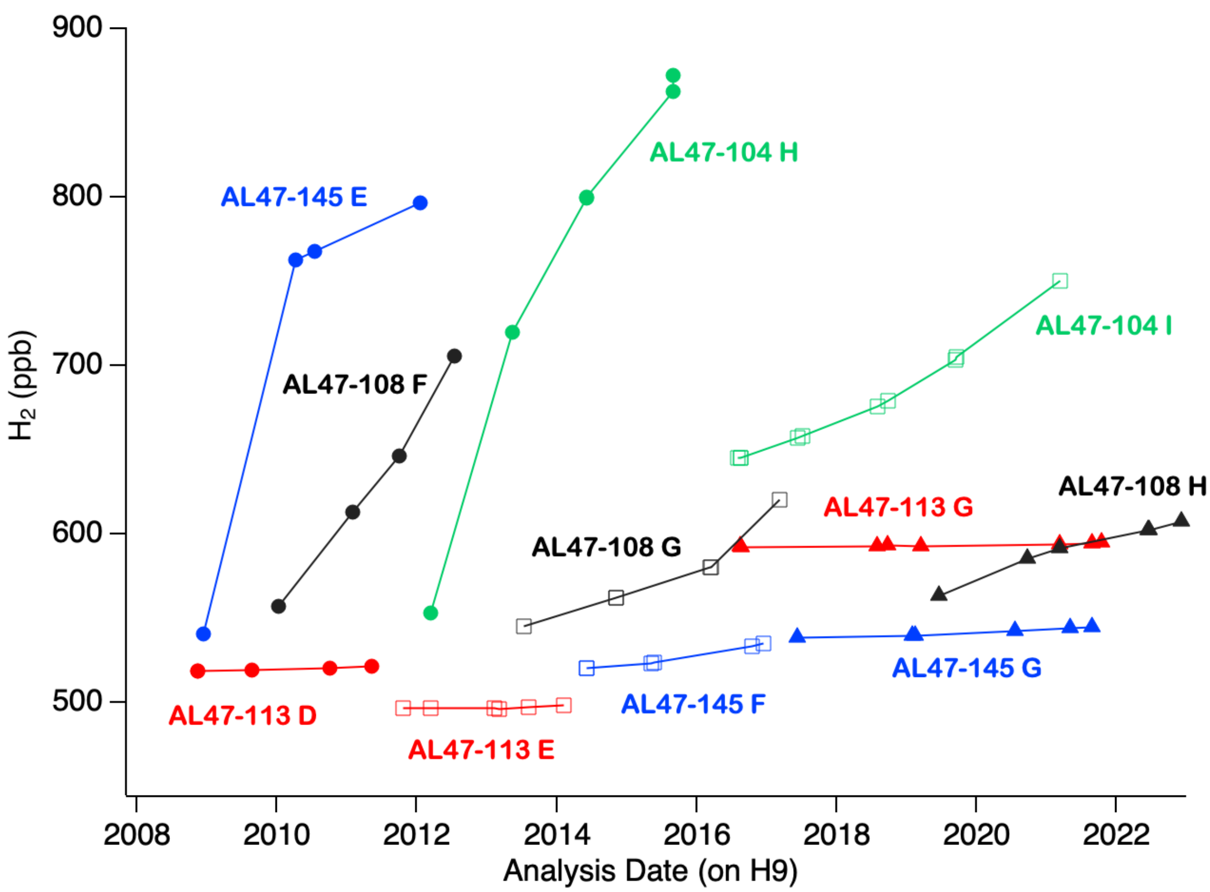


SI Figure 5: NOAA and MPI-BGC H_2 results for MPI-BGC GasLab led MENI tank air measurement round robin comparisons [Jordan and Damak, 2022]. NOAA measurement results are shown in blue. Asterix and open symbols show rejected results due to poor instrument performance or the use of an alternate calibration strategy respectively. All H9 tank air results for the period September 12-18, 2019 were biased high by a few ppbs. The reason is unknown at this point. Most MPI-BGC results (red symbols) are on their GC-PDD instrument, except the April 2020 results are from their GC-RGA instrument. a) Cylinder D232733 is a blind sample and is refilled with different air after each round robin analysis loop. b) Ambient H_2 cylinder D232733 (~565 ppb) and c) low H_2 cylinder D232717 (~ 335 ppb) have slightly increasing H_2. The NOAA and MPI-BGC H_2 results agree well for the ambient and blind H_2 MENI tanks (< 1 ppb difference).



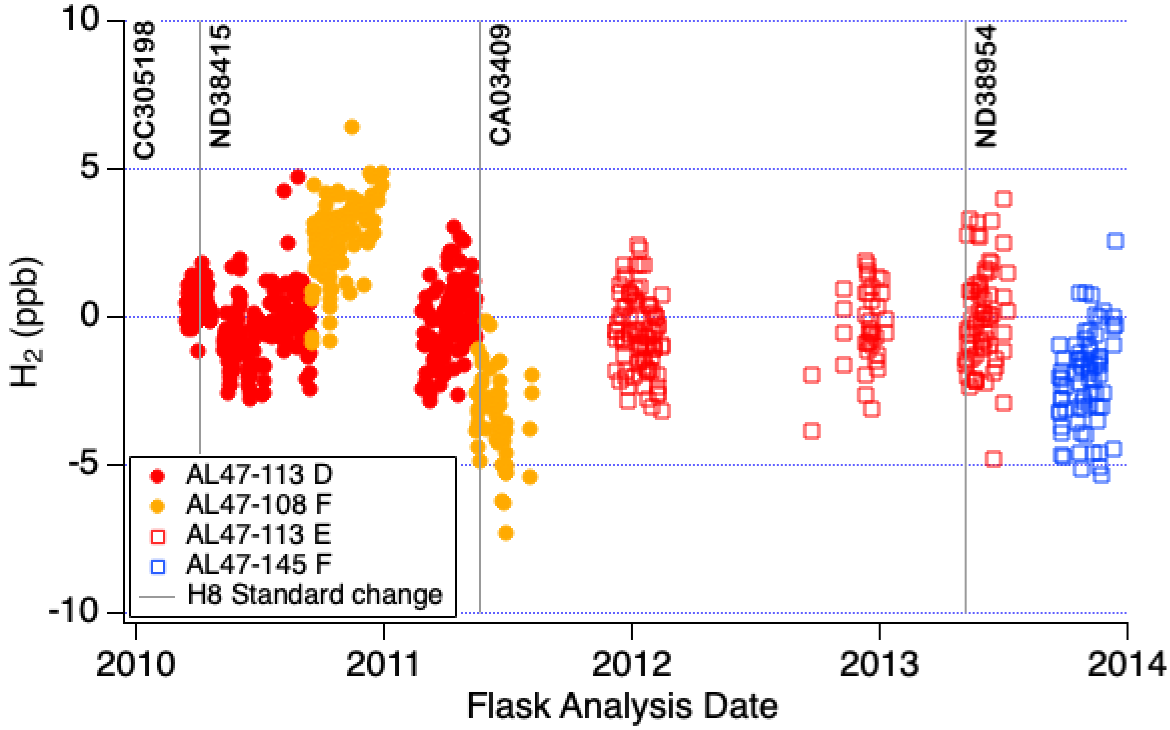


SI Figure 6. H_2 calibration histories of test air tanks 2008-2022. Each test air cylinder has a different color and different tank fills are shown with different symbols.

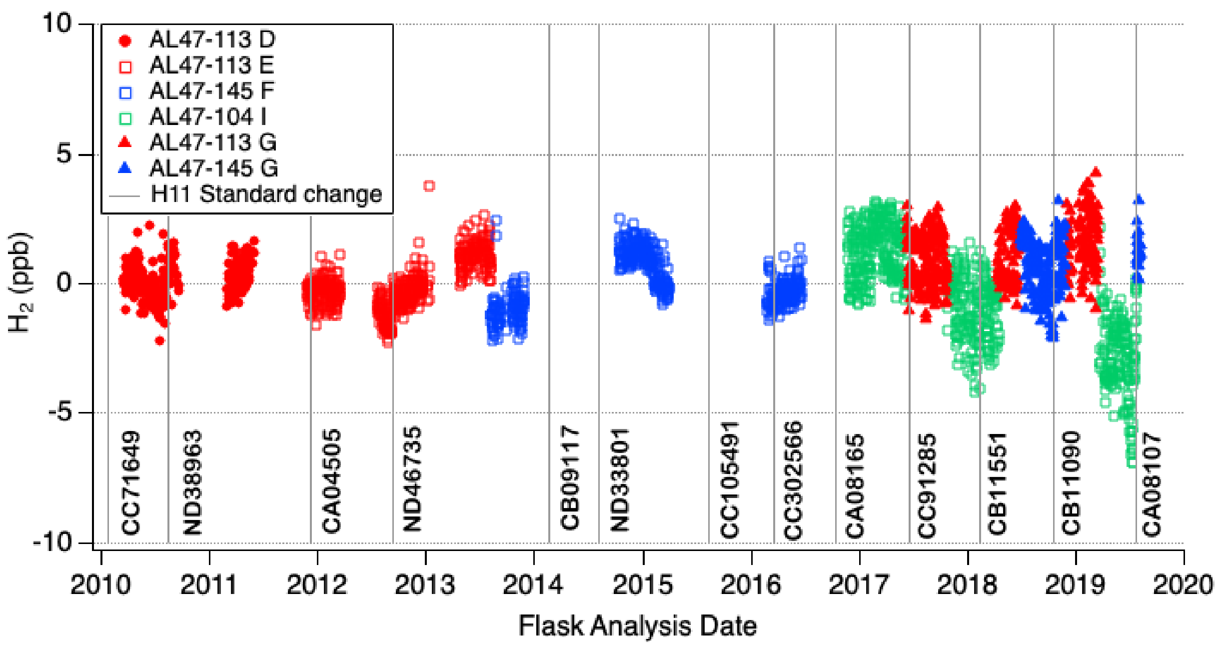


SI Figure 7. Test air (TST) flask analysis results : differences from test air tank time-dependent H_2 assignment: a) on H8, b) on H11 and c) on MAGICC-3.

a)

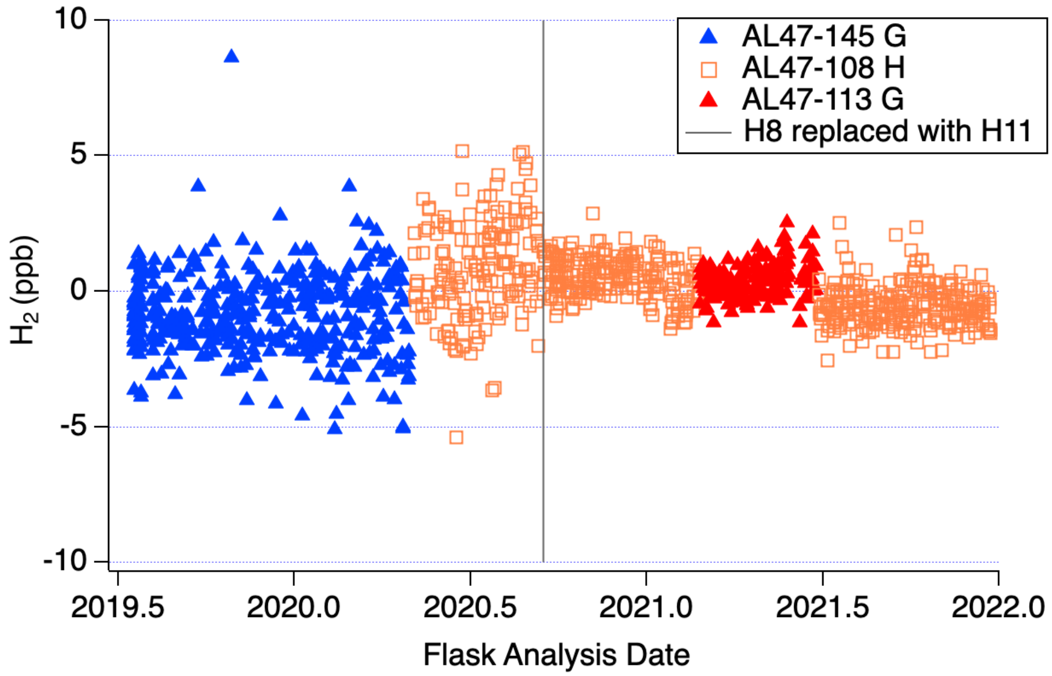


b)

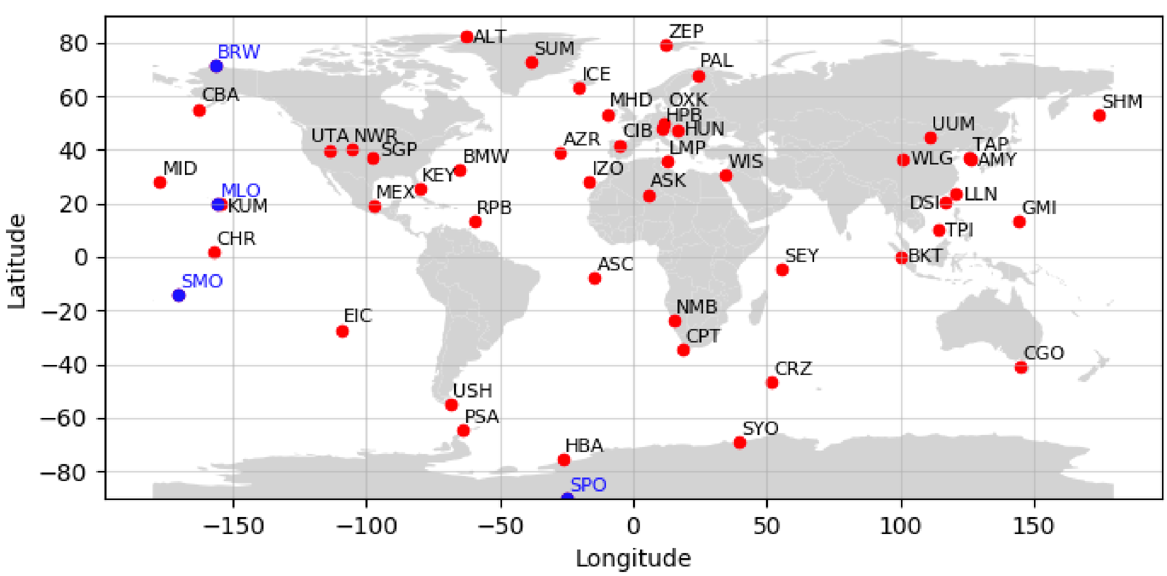


.

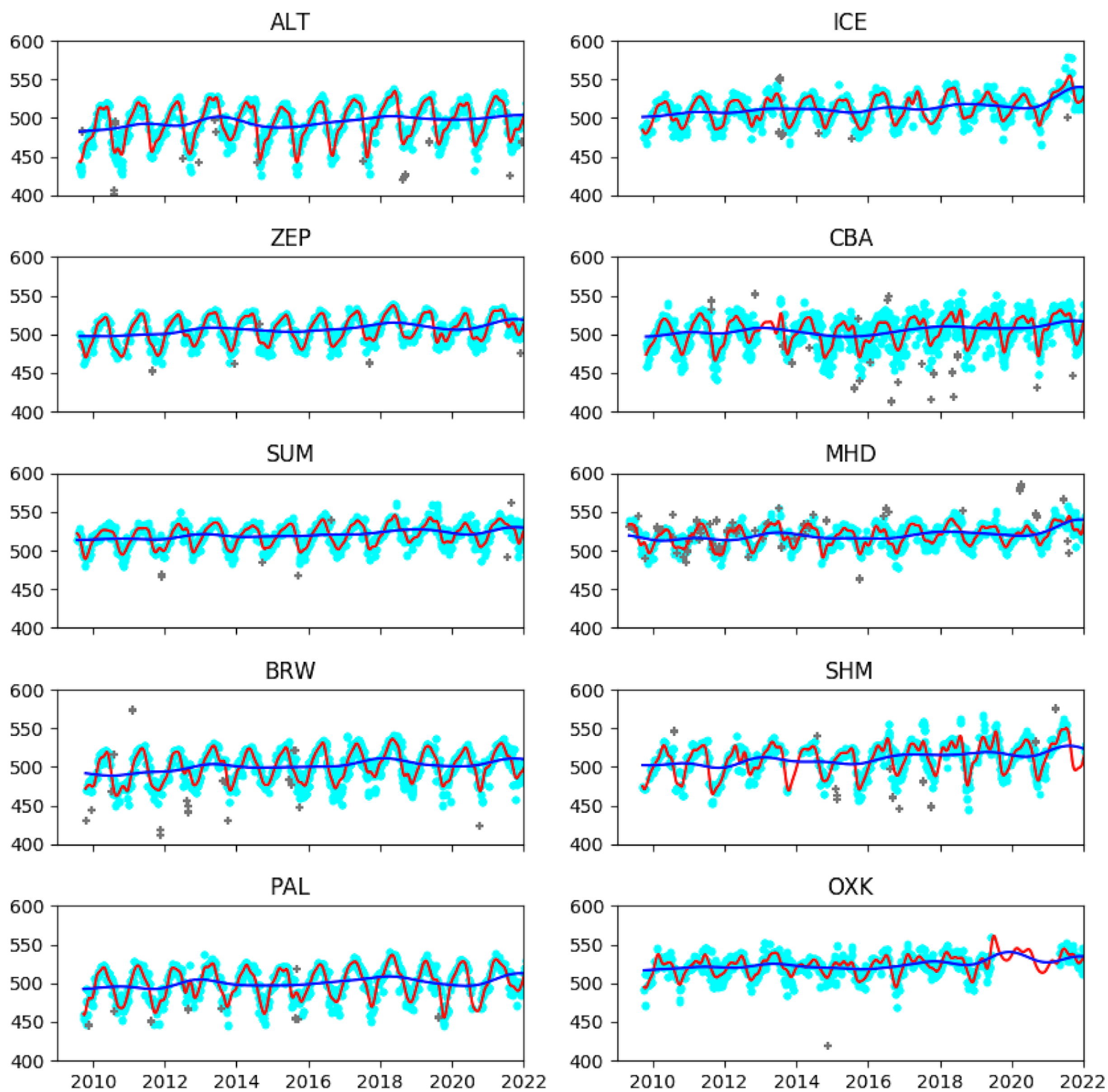
c)

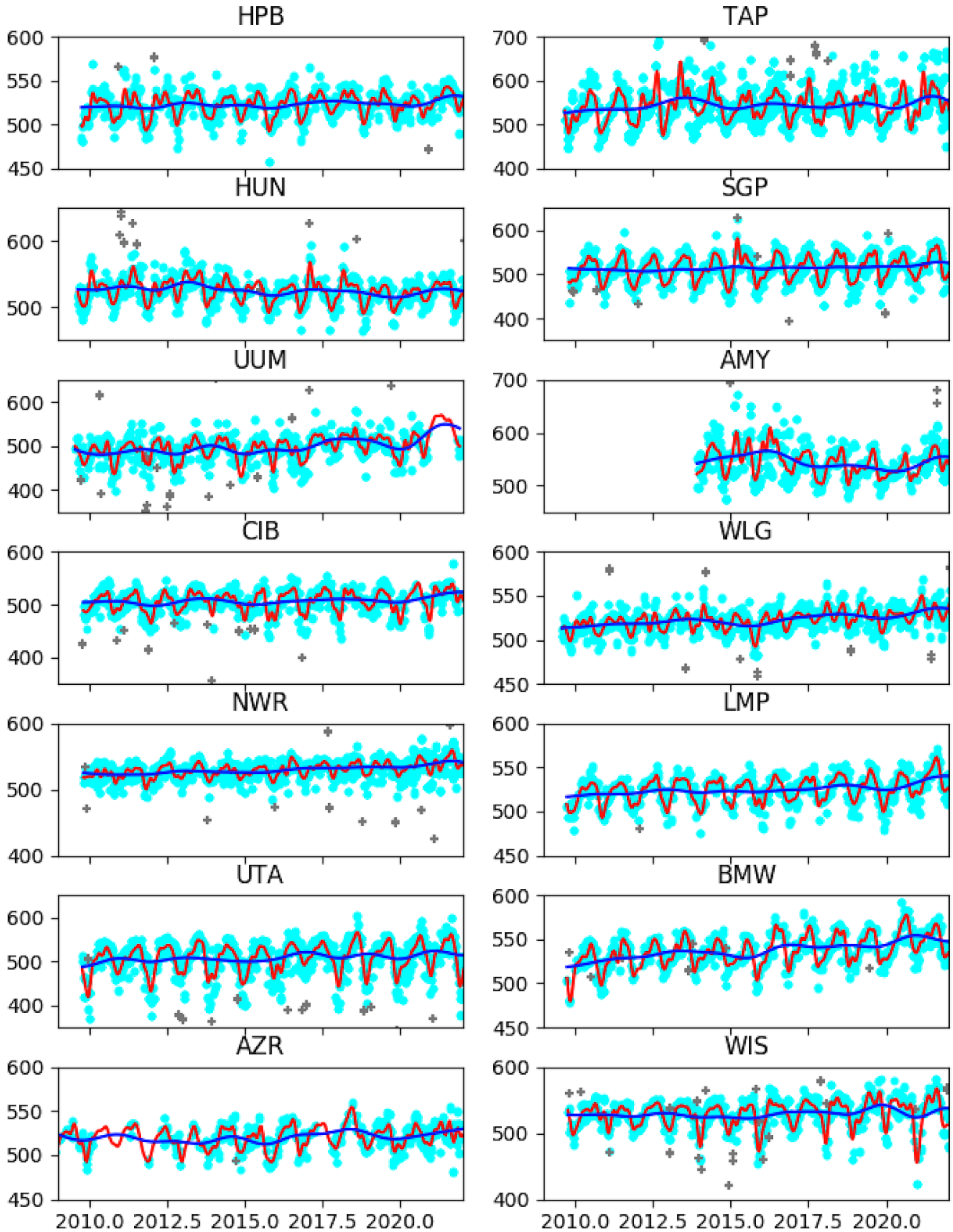


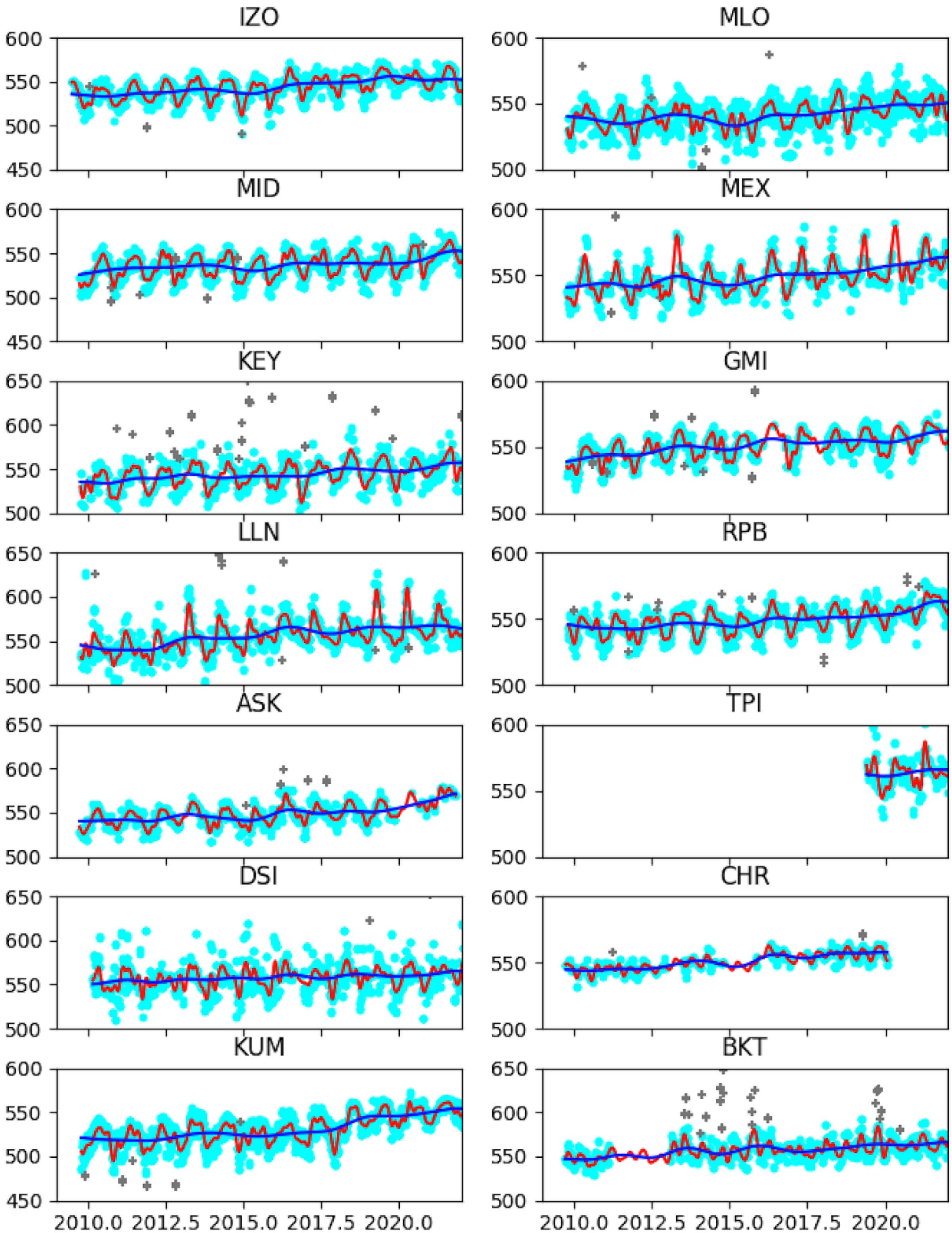
SI Figure 8: NOAA Global Cooperative Air Sampling Network site map (https://gml.noaa.gov/dv/site/). The four NOAA atmospheric baseline observatories are shown in blue.

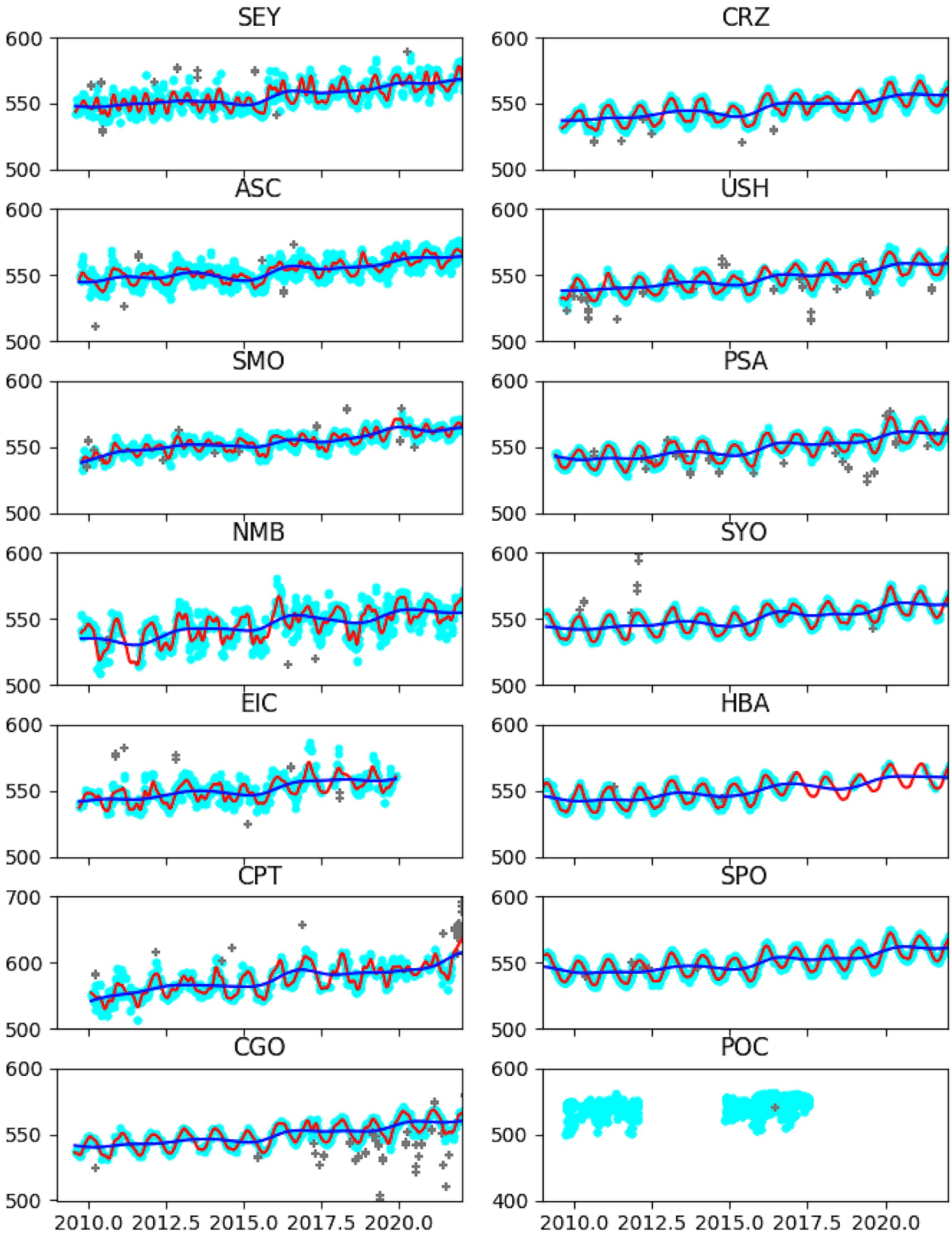


SI Figure 9: Discrete air H_2 mole fraction (in ppb) time series at 51 sites from the NOAA Global Cooperative Air Sampling Network. Data in light blue symbols are retained and data shown in gray crosses are deemed to be non-background. Rejected data are not shown but are present in the site data files. A curve fit python code is run for each site H_2 time series based on Thoning et al. [1989]. First the code optimizes parameters for a function made of a four-term harmonic and a cubic polynomial. The resulting residuals are then smoothed with a low-pass filter with a 667 day cutoff and are added to the polynomial part of the function to produce the “trend curve” shown as the dark blue line. The residuals are also smoothed with a low-pass filter with a 80 day cutoff and are added to the function to produce a “smooth curve”, a detrended and smoothed. The last plot shows all retained H_2 measurements from the Pacific Ocean Shipboard (POC).

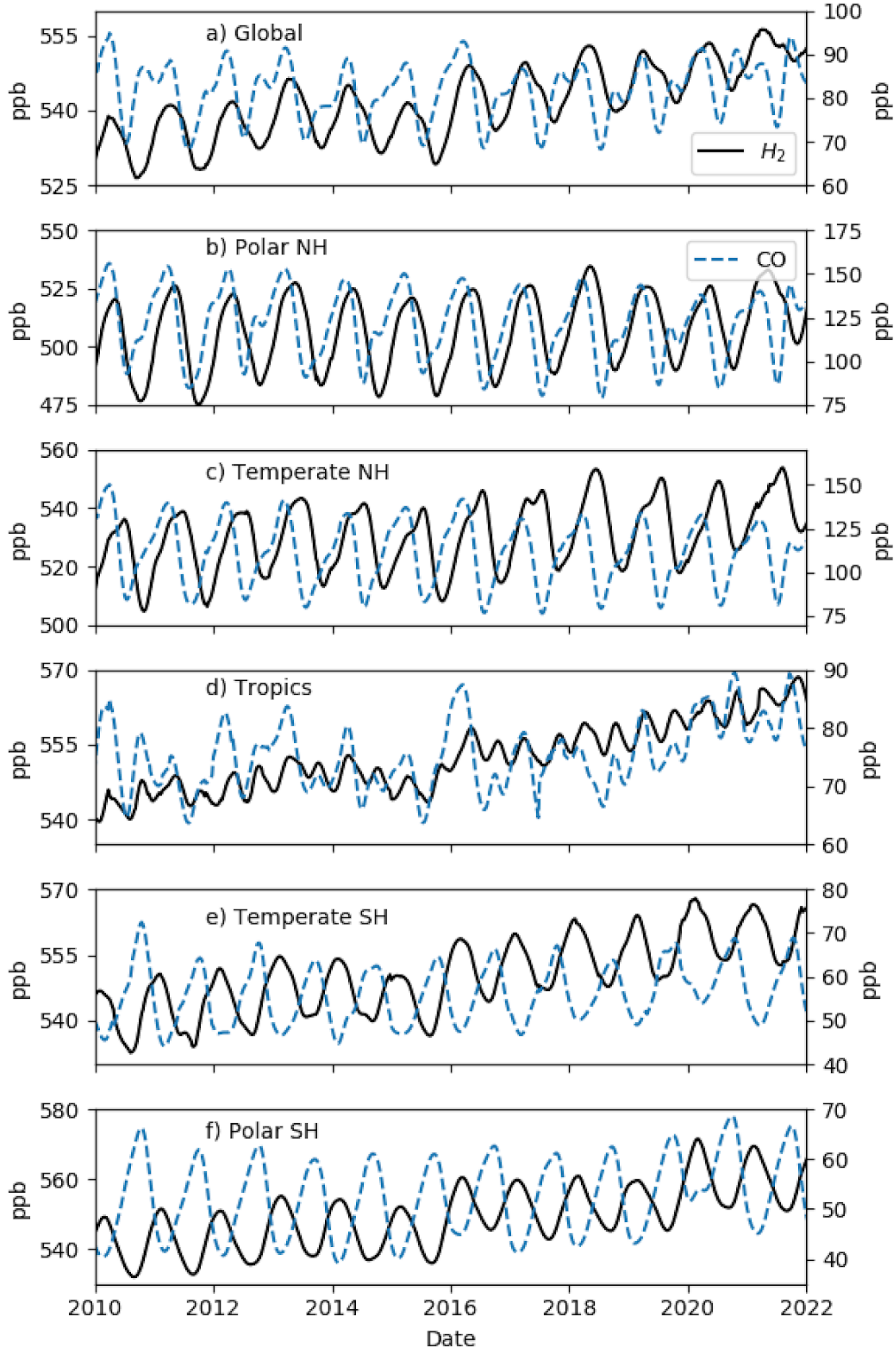



HPB
TAP
HUN
SGP
UUM
AMY
CIB
WLG
NWR
LMP
UTA
BMW
AZR
WIS
2010.0 2012.5 2015.0 2017.5 2020.0
2010.0 2012.5 2015.0 2017.5 2020.0

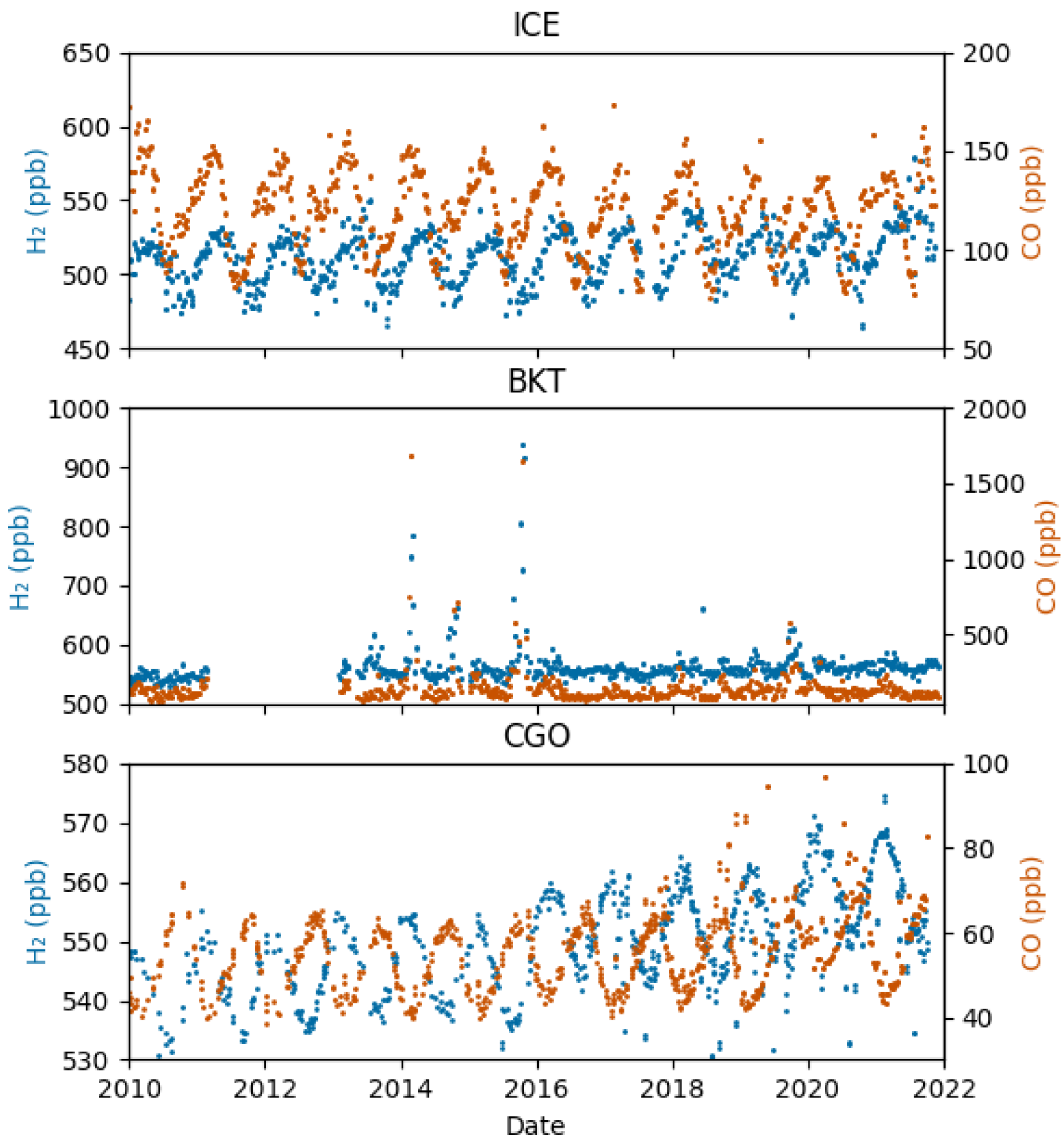

IZO
MLO
MID
MEX
KEY
GMI
LLN
RPB
ASK
TPI
DSI
CHR
KUM
BKT
2010.0 2012.5 2015.0 2017.5 2020.0
2010.0 2012.5 2015.0 2017.5 2020.0


SEY
CRZ
ASC
USH
SMO
PSA
NMB
SYO
EIC
HBA
CPT
SPO
CGO
POC
600
550
500
700
400
2010.0
2012.5
2015.0
2017.5
2020.0

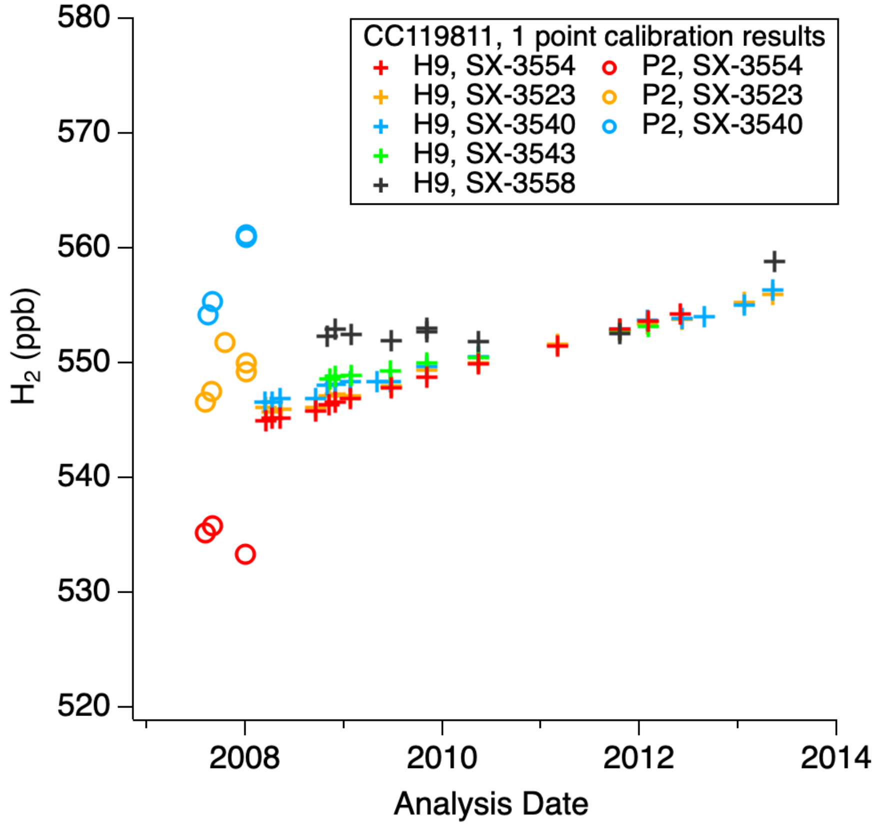
SI Figure 10: Marine boundary layer global mean and zonal mean H_2 (black, left side y axis) and CO (dashed blue line, right y axis) time series



SI Figure 11: NOAA H_2 and CO measurement times series for three Global Cooperative Air Sampling Network sites in Iceland (ICE: 63.3998°N, 20.2884° W, 118.00 masl), Indonesia (BKT: 0.202° S, 00.3180° E, 845.00 masl) and Tasmania, Australia (CGO: 40.683° S, 144.6900° E, 94.00 masl).



SI Figure 12: NOAA H2 secondary standard CC119811 results on Peak Labs instrument (P2) and on GC-HePDD H9 using one point calibration against one of the primary standards.



SI Figure 13: Early target tanks measurement records on different instruments using one point calibration. The working standard/reference tank ID for the measurements on RGA instruments is indicated in the legend.