A 60-year record of atmospheric carbon monoxide reconstructed from Greenland firn air: SUPPLEMENTARY MATERIAL

2. Sampling and Analytical Methods

We provide some further detail here on NOAA [CO] and [H_2] analyses, since up-to-date descriptions are not available elsewhere. [CO] was measured on a system designed to analyze a suite of trace gases in the same air sample. Sample flasks were first attached to an eight port sampling manifold fitted with all stainless steel components. The dead volume between flasks and a down-line transfer pump was then evacuated to < 0.13 mbar. A stream selection valve determines which flask is opened to the system and a transfer pump with a MFC regulates flow rates. Samples are cryogenically dried to -80 C before a second stream selection valve partitions air flow to the CO analyzer and three other instruments. Sample flow rates, stream selection valves and instrument functions are controlled by HP computer and custom software (http://www.esrl.noaa.gov/gmd/outreach/behind_the_scenes/measurementlab.html). Sample metadata and digitalized instrument response are archived. [CO] is calculated in units of nmol mol^{-1}.

Instruments (Reducing Gas Analyzer) from Trace Analytical Inc. were calibrated using six reference gases evenly spaced between 50 to 200 nmol mol^{-1} [CO]. Instrument response over this range was defined by a quadratic function. The VURF instruments were linear over the atmospheric range and therefore the calibration consisted of a single reference gas (~300 nmol mol^{-1}) and a blank (UHP zero air run through a trap containing Schutze reagent). Reference gases are tied to the WMO-2004 [CO] scale. The scale is based upon six sets of primary CO-in-air mixtures prepared using gravimetric techniques (Novelli et al., 1991) and has undergone revision to address drift with time (Novelli et al., 2003).

An informal comparison of [H_2] measurements in 2008 between the Max-Plank Institute for Biogeochemistry (MPI-BGC) and NOAA showed a difference of 9 nmol mol^{-1} over the range 470 to 560, with NOAA low. The MPI-BGC acts as the World Calibration Center for WMO/GAW H_2 reference gases.

Table S2 shows the data from the field procedural blank tests for the US and EU firn air systems (analyzed at NOAA). Pairs of flasks were filled using each system to test the system blanks. The average US system [CO] blank is -0.4 ppb, and the average EU system blank is + 0.5 ppb which are both smaller than the estimated NOAA measurement uncertainty (1.2 ppb). The US system [H_2] blank is 4.8 ppb (compared to measurement 1σ uncertainty of 4.0 ppb), and the EU system [H_2] blank is 6.1 ppb.

Figure S1 shows the [CO] and [H_2] comparisons of surface flasks filled through the firn air systems with NOAA flask measurements from the two closest monitoring stations during the same time period. All air samples shown (firn surface and NOAA) were taken using identical 2.5L glass flasks and were measured at NOAA. As can be seen, [CO] and [H_2] in surface air sampled with firn air systems at all sites are within the range expected

from the nearest NOAA flask collection sites, Summit (72.58˚N, 38.46˚W) and Alert (82.45˚N, 62.51˚W).

3.1 NEEM 2008 Firn Air Data

The UEA [CO] and [H2] data from the S4 borehole were excluded from the final combined data sets because we have decided that it is best to treat each borehole separately for the reconstructions (see main text), and the S4 borehole did not sample the oldest air and did not have enough supporting trace gas measurements to constrain the depth-diffusivity profiles (see Section 4.1) as well as in the EU and US boreholes. In the top ≈ 30 m, this borehole shows lower CO mole fractions than the other two boreholes (Fig. 3), which likely reflects the propagation of the seasonal [CO] minimum into the firn during the two weeks that elapsed between sampling the EU and US holes and S4. This process was likely enhanced due to the adjacent (≈ 15 m) open US borehole. Both UEA and NOAA [CO] data are nominally on the WMO 2004 [CO] scale, yet for unknown reasons the UEA data are 9 ppb lower than the NOAA data even for depths below 30 m, which should be unaffected by surface air. A single sample from the US borehole was collected with the UEA high-volume system from a 74.6 - 75.6 m depth range. Although the measured [CO] agrees with what would be expected based on other data from NEEM boreholes, this sample was not included in the final combined data set because the sampling system had to be positioned 1 m above the bottom of the borehole (rather than right at the bottom, as is the normal procedure) to achieve sufficient air flow.

Stony Brook [CO] data were excluded because of relatively larger analytical uncertainties; excluding the Stony Brook data does not strongly affect the finalized data set for the NEEM EU borehole (see Fig. S5). Heidelberg [H_2] data were excluded because the data set was clearly more noisy than the NOAA or CSIRO data, suggesting issues during flask storage or analysis.

In order to combine the different sets of firn measurements, they must first be placed on the same calibration scale. Calibration offsets for both [CO] and [H_2] between different laboratories are imperfectly characterized, however. For this reason we chose to use the NEEM data themselves to calculate the calibration offsets. Individual offsets were determined for each available depth level in each borehole as NOAA - Heidelberg and NOAA – CSIRO differences. The average of the NOAA - Heidelberg and NOAA - CSIRO offsets was then used to adjust the Heidelberg and CSIRO measurements to the WMO 2004 scale. All the sets of measurements from each borehole were then averaged (by depth level) to produce the finalized data sets shown in Figure 4. As discussed in the main text, we are not attempting to reconstruct a true atmospheric history of [H_2], but rather using [H_2] as a source of qualitative information about changes in the CO budget. For this reason, the inverse modeling for [H_2] was only done using the EU borehole data; these are the only finalized [H_2] data shown in Figure 4.

The error bars on the finalized NEEM data sets represent the overall data uncertainties and were determined as follows. For each depth level in each borehole, all the available

flasks (corrected to the WMO 2004 scale for [CO] and NOAA scale for [H_2]) were considered. Errors were first calculated for all depth levels with more than one flask sample as either standard deviation (3 or more flasks) or difference between the flask values (2 flasks). The average of these errors was then calculated for each species and each borehole. The final uncertainty for all depth levels with only a single flask is taken as this average. For depth levels with multiple flasks, the final uncertainty was taken as the larger of this average or the error calculated for that particular depth level.

3.3 Is CO well preserved in Greenland firn air?

As mentioned in the Introduction, the only firn air record of Northern Hemisphere [CO] published to-date indicated in-situ CO production in the Devon Ice Cap firn (Clark et al., 2007). The only reasonable substrate for such CO production, regardless of the exact mechanism, is trace organic material in the firn (e.g., Colussi and Hoffmann, 2003; Haan and Raynaud, 1998). Due to the small size of the Devon Ice Cap, the Devon firn air site is much closer to terrigenous sources of organic material than inland Greenland sites. The Devon site is also warmer (mean annual temperature -23 ˚C,) than any of our Greenland sites (-28.9˚C for NEEM (Buizert et al., 2012), -31.5˚C for NGRIP (Andersen et al., 2004) and -31˚C for Summit (Grootes et al., 1993)) and contains numerous melt layers (Clark et al., 2007), both of which may result in higher CO production rates in the ice depending on the mechanism.

Although inland Greenland sites are colder and cleaner with respect to trace organics, Haan and Raynaud (1998) still found evidence for in-situ CO production in ice below 155 m depth from the Summit region. Ice from shallower depths, however, yielded very stable [CO] values around 90 nmol mol^{-1} prior to 1862 AD and a gradual increase (as expected from the ramping up of anthropogenic emissions) to 100 nmol mol^{-1} by 1905 AD (Haan et al., 1996). Thus, published prior work allows both for the possibility that Greenland firn air [CO] is well preserved, as well as for the possibility that there is some alteration of the record.

To examine this issue, we first considered the possibility that the [CO] peaks observed in the lock-in zone at each site are artifacts of in-situ production from an ice layer (or layers) that is enriched in trace organics. Under this hypothesis, this organic – rich ice layer would have been deposited at all 3 sites. This hypothesis can be ruled out by comparing NEEM and NGRIP. At NEEM, the [CO] peak occurs at 70 m, which corresponds to an ice layer age of 225 years (GICC05modelext-NEEM-1 time scale, unpublished). At NGRIP, the [CO] peak occurs at 71.75 m, corresponding to an ice layer age of 270 years (GICC05 age scale, (Vinther et al., 2006)).

As stated in the main text, we also examined the issue of CO preservation at our sites by comparing firn data with LGGE-GIPSA firn model runs that used the available Arctic station [CO] data as input. To create the combined OGI-NOAA record for Barrow used in some of these runs, we used the time period of overlap between the two data sets (1988 – 1992) to determine the calibration scale offset (NOAA scale = OGI scale x 1.27) and to

place the OGI data on the NOAA calibration scale. No dependence of the scale offset on season (and therefore on absolute [CO] values) was observed. For all firn model runs presented in Figure 7, [CO] for dates prior to the start of atmospheric measurement time series is taken to equal our average reconstructed [CO] scenario (see Fig. 8). While the model-data agreement in Figure 7 is generally good in the upper firn, we note that our forward gas transport models do not do a perfect job with reproducing seasonal signals. For example, we observed some small mis-matches between data and models for the upper firn even for a gas with a modest seasonal cycle like CO_2 (Buizert et al., 2012). We estimate that for [CO], which has a very large seasonal cycle, seasonal effects are significant to a depth of ~40 m.

One possible explanation for the 0 - 5 nmol mol^{-1} elevation of data over the model runs is that the firn values are weighted toward the winter surface mixing ratios. [CO] has a very large seasonal cycle over Greenland, with peak values in the winter. Figure S3 shows a comparison of the [CO] seasonal cycle to the seasonal variability in surface pressure and wind from the Summit and Humboldt automated weather stations (AWS; Fig. 2). These two AWSs cover the full geographic range of our firn air sites. The Humboldt AWS was chosen instead of NEEM because of its longer and more continuous record (NEEM record only starts in 2007 and does not include most of the winters). As can be seen from Figure S3, both surface pressure variability and wind speed correlate with [CO]. Both would be expected to enhance gas exchange between the surface and the upper firn (e.g., Colbeck, 1989; Schwander, 1989; Sowers et al., 1992), resulting in an effectively higher gas diffusivity in the winter.

It is also possible that this winter weighting effect is slightly different between our sites and explains some of the small differences between sites evident in Figures 6 and 7. For example, NEEM is closer to the Humboldt weather station and therefore likely experiences higher winds than Summit in the winter. Also, as discussed in the main text, Summit appears to have a reduced seasonal cycle for [CO] as compared to other sites, with lower [CO] in the winter. NEEM is approximately mid-way between Summit and Alert and would likely see higher winter [CO] than Summit. Thus, the winter weighting effect may result in a slightly higher [CO] recorded in the firn at NEEM than at Summit and NGRIP, consistent with what Figure 6 shows.

One further effect that can influence the data – model comparisons is a likely small drift in the NOAA [CO] calibration scale over time. This drift is currently being investigated at NOAA ESRL with the ultimate objective of correcting all the affected data. For atmospheric data used in this study, values from NOAA measurements prior to 2000 are likely too low by 2 – 4 ppb. This would affect the data-model comparisons in Figure 7, making the model curves appear too low for the deepest compared data points. The calibration scale shift is estimated to have happened gradually between 2000 and 2004, thus the NGRIP firn data may also be biased low by 2 - 4 ppb. Summit firn data and NEEM firn data are unaffected by this.

4.1 Forward modeling

There are some important differences in the ways the LGGE-GIPSA (Witrant et al., 2012) and the INSTAAR (Buizert et al., 2012) forward models handle gas transport in the firn. The INSTAAR model includes molecular diffusion (Fick's Law; different for different gases), eddy diffusion (also parameterized as Fick's law, but same for all gases), and a downward advection term calculated based on mass conservation. Gravitational settling is explicitly parameterized in the INSTAAR model by including a $\Delta Mg/RT$ term in the gas transport equations (see Buizert et al. (2012) for full gas transport equations). The model physics are uniform throughout the firn column, and both molecular and eddy diffusivities are manually tuned at each depth level to give the optimal overall fit for a suite of reference gas species (Buizert et al., 2012). The INSTAAR model has explicit time stepping and a depth resolution of 1 m down to 59.5 m, and 0.25m below that level.

The LGGE-GIPSA model (Witrant et al., 2012) treats molecular diffusion in a similar way as the INSTAAR model. It also takes into account an eddy diffusion term in the upper firn which represents fast ("convective") exchanges with the atmosphere but its magnitude and vertical structure are different in the two models (Buizert et al., 2012). The LGGE-GIPSA model does not take into account a similar eddy diffusion term aiming at representing dispersive transport in the lock-in zone. Because the CO diffusion coefficient in free air is significantly higher than that of CO_2, the relative weights of molecular and eddy diffusion coefficients in the two models can potentially affect the consistency of their results. The LGGE-GIPSA model representation of gravitational settling is based on a quasi-steady-state approximation of Darcy's law, which allows to predict the location of the lock-in depth. As the LGGE-GIPSA and INSTAAR models lead to very similar results for $\delta^{15}N$ of N_2, the different formulation of gravitational setting in the two models does not have a significant impact on the results for CO. The LGGE-GIPSA model representation of advective transport and bubble closure in the lock-in zone follows the approach of Rommelaere et al. (1997). Although these processes are formulated differently in the INSTAAR model, a specific test of these processes in Buizert et al. (2012) (synthetic scenario IV) showed that the two models produce nearly equivalent results. The LGGE-GIPSA model uses implicit time stepping and a depth resolution of 0.2 m throughout the firn.

The tuning of diffusivities in each model largely compensates for the differences in model physics, as demonstrated by the overall similar performance of the models in a recent intercomparison (Buizert et al., 2012). Further, the fact that our reconstructed [CO] time trend scenarios can be successfully used in both models to reproduce the firn [CO] data (see Section 4.2 and Figure S5) shows that the two models perform similarly for CO.

Neither model includes thermal fractionation (e.g., Severinghaus et al., 2001) as this process is not expected to significantly affect CO. Both models use monthly-averaged atmospheric histories for model – data comparisons such as those shown in Figure 7. This low temporal resolution of the input histories could affect the model-data fit in the upper firn to a depth of at least 40 m. Further, neither model explicitly takes into account synoptic surface pressure variations. The average effective increase in diffusivity due to

this process is accounted for by diffusivity tuning, however this could still affect gases with large seasonal cycles (such as CO), as discussed in Section 3.3 above.

4.2 Inverse modeling and [CO] scenario generation

As discussed in the main text, atmospheric trend reconstruction from firn data is a multiple-solution inverse problem and efforts have been made recently to improve the definition of the optimal solution (Witrant and Martinerie, 2013). The smoothness of a scenario reconstructed from firn data results from both the physical and mathematical nature of the problem:

(1) The deep firn imposes a large amount of smoothing on the atmospheric signal, with the degree of smoothing increasing with depth. For example, at the start of the lock-in zone at NEEM, the width at half-height for the CO gas age distribution is already ~8 years, and it is ~35 years at the deepest sampled level (Figure S2).

(2) The mathematical inverse problem best estimate smoothness is related to what the model considers as noise and as signal in the firn records. This smoothness thus depends on the quality of the data, but also on the model quality (capacity to fit the data well) and assumptions (e.g. the atmospheric trend scenario is the same at all modeled sites).

Atmospheric history scenarios with low smoothing (a large amount of temporal variability) can be generated that match the firn data better when run through our two forward models than smoother (more robust) scenarios. We cannot rule such scenarios out for CO, as a large amount of interannual variability is likely given the relatively short lifetime of CO. However, the firn does not contain information of sufficient temporal resolution to validate the higher frequency components of such scenarios. The two-step scenario selection approach in the main manuscript thus aims to better explore the interplay of non-modelled uncertainties and mathematical under-determination of the problem by testing reasonably smooth scenarios through all available datasets and models.

The new development of a robustness-oriented solution (Witrant and Martinerie, 2013) leads to smooth scenarios shown in blue and red on Fig. S4. The use of different firn models (LGGE-GIPSA or INSTAAR) and/or different combination of sites still leads to a wide range of [CO] peak date estimates. Smoother scenarios tend to fit the firn data less well, and NEEM-EU+NEEM-US constrained scenarios still have difficulty matching the North GRIP data point at 42.6 m depth. Importantly, the model-calculated uncertainty on the new robust scenarios constrained with three sites (dashed green lines on Figure S4) are similar to the range of scenarios from our two-step approach. This confirms the overall magnitude of the uncertainties for the original set of reconstructed [CO] scenarios.

References

Andersen, K. K., Azuma, N., Barnola, J. M., Bigler, M., Biscaye, P., Caillon, N., Chappellaz, J., Clausen, H. B., Dahl-Jensen, D., Fischer, H., Fluckiger, J., Fritzsche, D., Fujii, Y., Goto-Azuma, K., Gronvold, K., Gundestrup, N. S., Hansson, M., Huber, C., Hvidberg, C. S., Johnsen, S. J., Jonsell, U., Jouzel, J., Kipfstuhl, S., Landais, A., Leuenberger, M., Lorrain, R., Masson-Delmotte, V., Miller, H., Motoyama, H., Narita, H., Popp, T., Rasmussen, S. O., Raynaud, D., Rothlisberger, R., Ruth, U., Samyn, D., Schwander, J., Shoji, H., Siggard-Andersen, M. L., Steffensen, J. P., Stocker, T., Sveinbjornsdottir, A. E., Svensson, A., Takata, M., Tison, J. L., Thorsteinsson, T., Watanabe, O., Wilhelms, F., White, J. W. C., and Project, N. G. I. C.: High-resolution record of Northern Hemisphere climate extending into the last interglacial period, Nature, 431, 147-151, 2004.

Bergamaschi, P., Hein, R., Heimann, M., and Crutzen, P. J.: Inverse modeling of the global CO cycle 1. Inversion of CO mixing ratios, J. Geophys. Res. - Atmos., 105, 1909-1927, 2000.

Buizert, C., Martinerie, P., Petrenko, V. V., Severinghaus, J., Trudinger, C., Witrant, E., Rosen, J. L., Orsi, A. J., Rubino, M., Etheridge, D. M., Steele, L. P., Hogan, C., Laube, J. C., Sturges, W. T., Levchenko, V., Smith, A. M., Levin, I., Conway, T. J., Dlugkencky, E. J., Lang, P. M., Kawamura, K., Jenk, T. M., White, J. W. C., Sowers, T., Schwander, J., and Blunier, T.: Gas transport in firn: multiple-tracer characterisation and model intercomparison for NEEM, Northern Greenland, Atmos. Chem. Phys., 12, 4259-4277, 2012.

Clark, I. D., Henderson, L., Chappellaz, J., Fisher, D., Koerner, R., Worthy, D. E. J., Kotzer, T., Norman, A. L., and Barnola, J. M.: CO_2 isotopes as tracers of firn air diffusion and age in an Arctic ice cap with summer melting, Devon Island, Canada, J.Geophys. Res. - Atmos., 112, D01301, doi:10.1029/2006JD007471, 2007.

Colbeck, S. C.: Air Movement in Snow Due to Wind Pumping. J. Glaciol., 35, 209-213, 1989.

Colussi, A. J., and Hoffmann, M. R.: In situ photolysis of deep ice core contaminants by Cerenkov radiation of cosmic origin, Geophys. Res. Lett., 30 (4)**,** 1195, doi:10.1029/2002GL016112, 2003.

Duncan, B. N., Logan, J. A., Bey, I., Megretskaia, I. A., Yantosca, R. M., Novelli, P. C., Jones, N. B., and Rinsland, C. P.: Global budget of CO, 1988-1997: Source estimates and validation with a global model, J. Geophys. Res. - Atmos., 112, D22301, doi:10.1029/2007JD008459, 2007.

Grootes, P. M., Stuiver, M., White, J. W. C., Johnsen, S., and Jouzel, J.: Comparison of Oxygen-Isotope Records from the GISP2 and GRIP Greenland Ice Cores, Nature 366, 552-554, 1993.

Haan, D., Martinerie, P., and Raynaud, D.: Ice core data of atmospheric carbon monoxide over Antarctica and Greenland during the last 200 years, Geophys. Res. Lett., 23, 2235-2238, 1996.

Haan, D., and Raynaud, D.: Ice core record of CO variations during the last two millennia: atmospheric implications and chemical interactions within the Greenland ice, Tellus B, 50, 253-262, 1998.

Novelli, P. C., Elkins, J. W., and Steele, L. P.: The Development and Evaluation of a Gravimetric Reference Scale for Measurements of Atmospheric Carbon Monoxide, J. Geophys. Res. - Atmos., 96, 13109-13121, 1991.

Novelli, P. C., Masarie, K. A., and Lang, P. M.: Distributions and recent changes of carbon monoxide in the lower troposphere, J. Geophys. Res. - Atmos., 103, 19015-19033, 1998.

Novelli, P. C., Masarie, K. A., Lang, P. M., Hall, B. D., Myers, R. C., and Elkins, J. W.: Reanalysis of tropospheric CO trends: Effects of the 1997-1998 wildfires, J. Geophys. Res. - Atmos., 108, (D15), 4464, doi:10.1029/2002JD003031, 2003.

Pieterse, G., Krol, M.C., Batenburg, A.M., Steele, L.P., Krummel, P.B., Langenfelds, R.L., Rockmann, T.: Global modelling of H-2 mixing ratios and isotopic compositions with the TM5 model. Atmos. Chem. Phys., 11, 7001-7026, 2011.

Price, H., Jaegle, L., Rice, A., Quay, P., Novelli, P. C., and Gammon, R.: Global budget of molecular hydrogen and its deuterium content: Constraints from ground station, cruise, and aircraft observations, J. Geophys. Res. - Atmos., 112, D22108, doi:10.1029/2006JD008152, 2007.

Rommelaere, V., Arnaud, L., and Barnola, J. M.: Reconstructing recent atmospheric trace gas concentrations from polar firn and bubbly ice data by inverse methods. J. Geophys. Res. - Atmos., 102, 30069-30083, 1997.

Schwander, J.: The transformation of snow to ice and the occlusion of gases, in: The Environmental Record in Glaciers and Ice Sheets (H. Oeschger, and C. Langway, Eds.), pp. 53-67. John Wiley, New York, 1989.

Severinghaus, J. P., Grachev, A., and Battle, M.: Thermal fractionation of air in polar firn by seasonal temperature gradients, Geochem. Geophy. Geosy., 2, 2000GC000146, 2001.

Sowers, T., Bender, M., Raynaud, D., and Korotkevich, Y. S.: Delta-N-15 of N_2 in Air Trapped in Polar Ice - a Tracer of Gas Transport in the Firn and a Possible Constraint on Ice Age - Gas Age Differences, J. Geophys. Res. - Atmos., 97, 15683-15697, 1992.

Vinther, B. M., Clausen, H. B., Johnsen, S. J., Rasmussen, S. O., Andersen, K. K., Buchardt, S. L., Dahl-Jensen, D., Seierstad, I. K., Siggaard-Andersen, M. L., Steffensen, J. P., Svensson, A., Olsen, J., and Heinemeier, J.: A synchronized dating of three Greenland ice cores throughout the Holocene, J. Geophys. Res. - Atmos., 111, D13102, doi: 10.1029/2005jd006921, 2006.

Wang, Z., Chappellaz, J., Martinerie, P., Park, K., Petrenko, V. V., Witrant, E., Emmons, L. K., Blunier, T., Brenninkmeijer, C. A. M., and Mak, J. E.: The isotopic record of Northern Hemisphere atmospheric carbon monoxide since 1950: implications for the CO budget. Atmos. Chem. Phys., 12, 4365-4377, 2012.

Witrant, E., Martinerie, P., Hogan, C., Laube, J. C., Kawamura, K., Capron, E., Montzka, S. A., Dlugkencky, E. J., Etheridge, D., Blunier, T., and Sturges, W. T.: A new multi-gas constrained model of trace gas non-homogeneous transport in firn: evaluation and behavior at eleven polar sites, Atmos. Chem. Phys., 12, 11465-11483, doi:10.5194/acp-12-11465-2012, 2012.

Witrant, E. and Martinerie, P.: "Input Estimation from Sparse Measurements in LPV Systems and Isotopic Ratios in Polar Firns", Proc. of the 5th IFAC Symposium on System Structure and Control, pp 659-664, doi:10.3182/20130204-3-FR-2033.00201, 2013.

Borehole	Sampling Period	Gas	Laboratory
NEEM EU	July 2008	CO	NOAA, CSIRO, Heidelberg, Stony Brook*
		H_2	NOAA, CSIRO
NEEM US	July 2008	CO	NOAA, Heidelberg
		H_2	NOAA, Heidelberg*
NEEM S4	July 2008	CO	UEA*
		H_2	UEA*
Summit	May - June 2006	CO	NOAA
NGRIP	May - June 2001	CO	NOAA

Table S1. A summary of the Greenland firn air CO and H_2 measurements presented in this paper. Data sets that were not included in the final atmospheric history reconstructions are marked by * next to the name of the measurement laboratory.

	[CO], nmol mol^{-1}	[H_2], nmol mol^{-1}
Direct flask fill	4.0	800.1
Direct flask fill	4.5	798.4
Through US system	3.8	804.7
Through US system	3.8	803.5
Through EU system	4.5	809.2
Through EU system	5.1	801.5

Table S2. Results of the procedural blank tests conducted during NEEM firn air sampling; measurements were performed at NOAA.

NEEM EU 2008			NEEM US 2008			Summit 2006			NGRIP 2001	
Depth, m	**[CO]**		**Depth, m**	**[CO]**		**Depth, m**	**[CO]**		**Depth, m**	**[CO]**
0.00	88.3		0.00	93.9		0.00	129.4		0.30	117.2
4.90	106.3		2.85	94.2		15.05	143.9		2.40	128.7
10.10	134.3		5.23	128.5		25.00	139.2		4.97	132.3
14.80	133.7		9.83	129.8		29.96	133.8		7.50	136.2
19.75	135.9		19.30	134.8		39.92	133.8		10.05	141.1
27.54	136.8		34.70	132.8		50.00	134.8		14.95	142.4
34.72	134.8		49.70	130.5		58.00	136.3		20.00	138.9
50.00	132.8		57.47	137.2		63.00	136.1		27.43	135.8
57.40	134.3		59.90	137.3		65.89	137.3		34.63	134.6
59.90	135.9		62.00	136.3		67.95	136.8		42.57	136.5
61.95	135.2		64.03	142.2		70.13	138.0		55.14	141.0
63.85	140.4		65.50	144.5		72.17	144.9		59.35	141.3
65.75	146.1		66.90	151.0		74.30	150.1		62.30	142.2
68.05	154.5		68.30	154.8		76.00	153.2		65.02	142.7
70.05	157.0		69.80	155.8		78.00	151.2		66.99	145.0
72.00	152.4		71.40	152.0		79.94	148.8		69.04	150.3
74.08	146.7		72.85	145.9					71.75	153.0
75.90	141.7		73.80	143.6					74.30	148.2
77.75	135.9		75.60	138.5					76.70	143.7
									77.68	138.6

Table S3. Finalized, depth-averaged firn air [CO] data (in nmol mol^{-1}) from all boreholes that were used for atmospheric [CO] reconstructions. All data are on the WMO 2004 [CO] scale.

	Alert - Summit	Alert - Barrow	Alert - Ny Alesund
Period of overlap	Jul 1997 - Dec 2008 (large gaps)	Apr 1992 - Dec 2008	Feb 1994 - Dec 2008
Mean offset	0.2	-1.3	-1.8
January	12.7	-0.3	0.0
February	4.7	-2.8	1.1
March	5.8	-0.7	0.3
April	9.2	0.3	0.7
May	7.8	-0.1	1.8
June	-2.1	-2.2	-0.9
July	-8.7	-2.7	-1.4
August	-13.7	-3.5	-4.7
September	-6.0	-4.0	-3.5
October	-1.0	-1.3	-3.0
November	-1.7	0.6	-5.6
December	3.4	2.0	-6.6

Table S4. Comparison of [CO] offsets between NOAA Arctic flask sampling stations. The overall mean offset and the average offset for each month are calculated based on monthly data downloaded from ftp://ftp.cmdl.noaa.gov/ccg.

	CO	H_2
TOTAL SOURCE, Tg /yr	2236-2489	73
Source or sink partitioning, % of total		
SOURCES		
Fossil fuels	~20%	~25%
Biofuels	~8%	~6%
Biomass burning	~22%	~14%
N2 fixation in ocean	0%	~8%
Methane oxidation	~35%	~34%
Biogenic NMHC oxidation	~16%	~13%
SINKS		
OH	~90%	~25%
Soils	~10%	~75%

Table S5. Estimates of global sources and sinks of CO and H_2. CO budget is as in Duncan et al. (2007). OH and soil sinks for CO are from Bergamaschi et al. (2000) as listed in Table 2 of Duncan et al. (2007). H_2 budget is as in Price et al. (2007). The fossil fuel, biofuel and biomass burning terms also include CO and H_2 produced from oxidation of VOCs released from these combustion sources. We note that a more recent assessment of the H_2 budget has been presented in Pieterse et al. (2011); with very similar estimates to the Price et al. (2007) budget shown here.

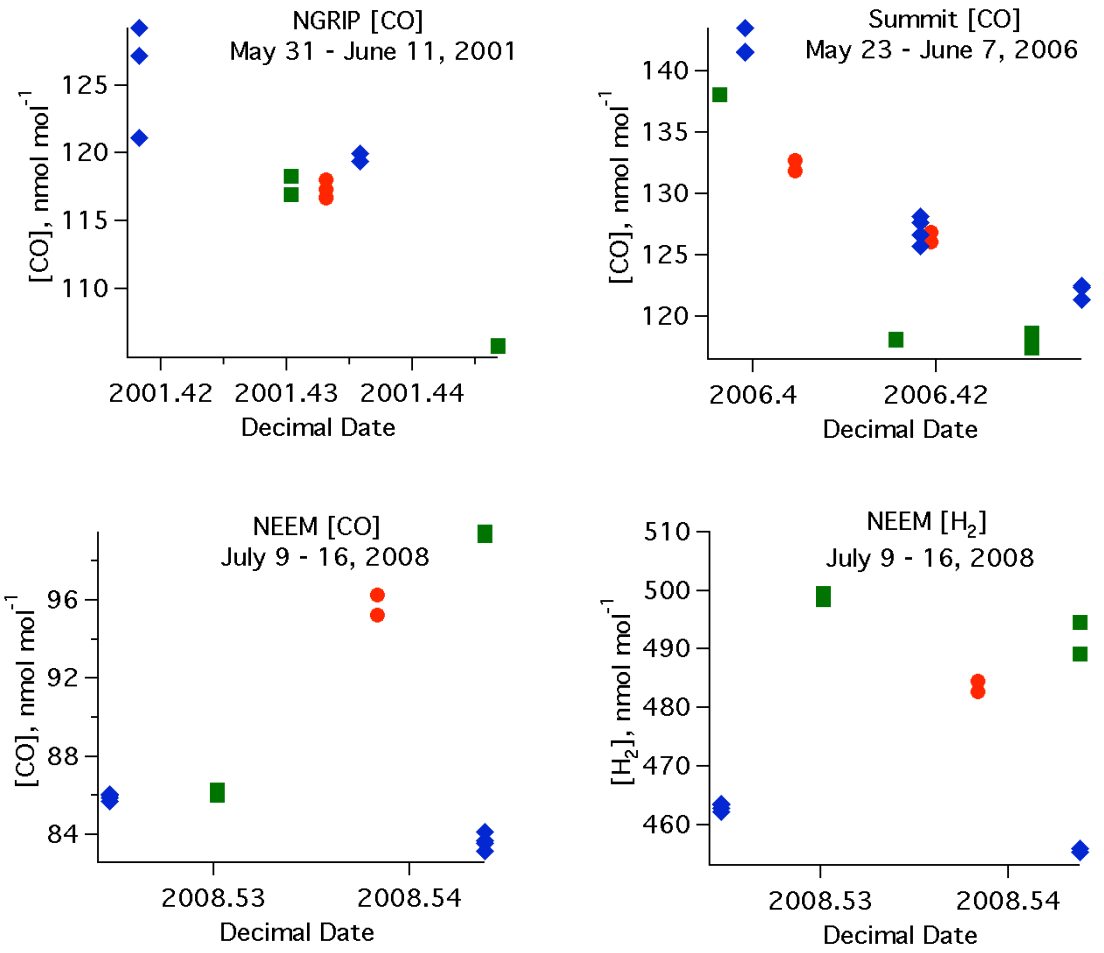

Figure S1. [CO] and [H_2] comparisons of surface flasks filled through the firn air systems with NOAA flask measurements from the two closest monitoring stations during the same time period. Red circles: surface air samples from firn sampling sites; green squares: NOAA Summit data; blue diamonds: NOAA Alert data. All [CO] data are on WMO 2004 scale, all [H_2] data are on NOAA scale. Note that large short-term fluctuations are not uncommon for these gases (especially for CO).

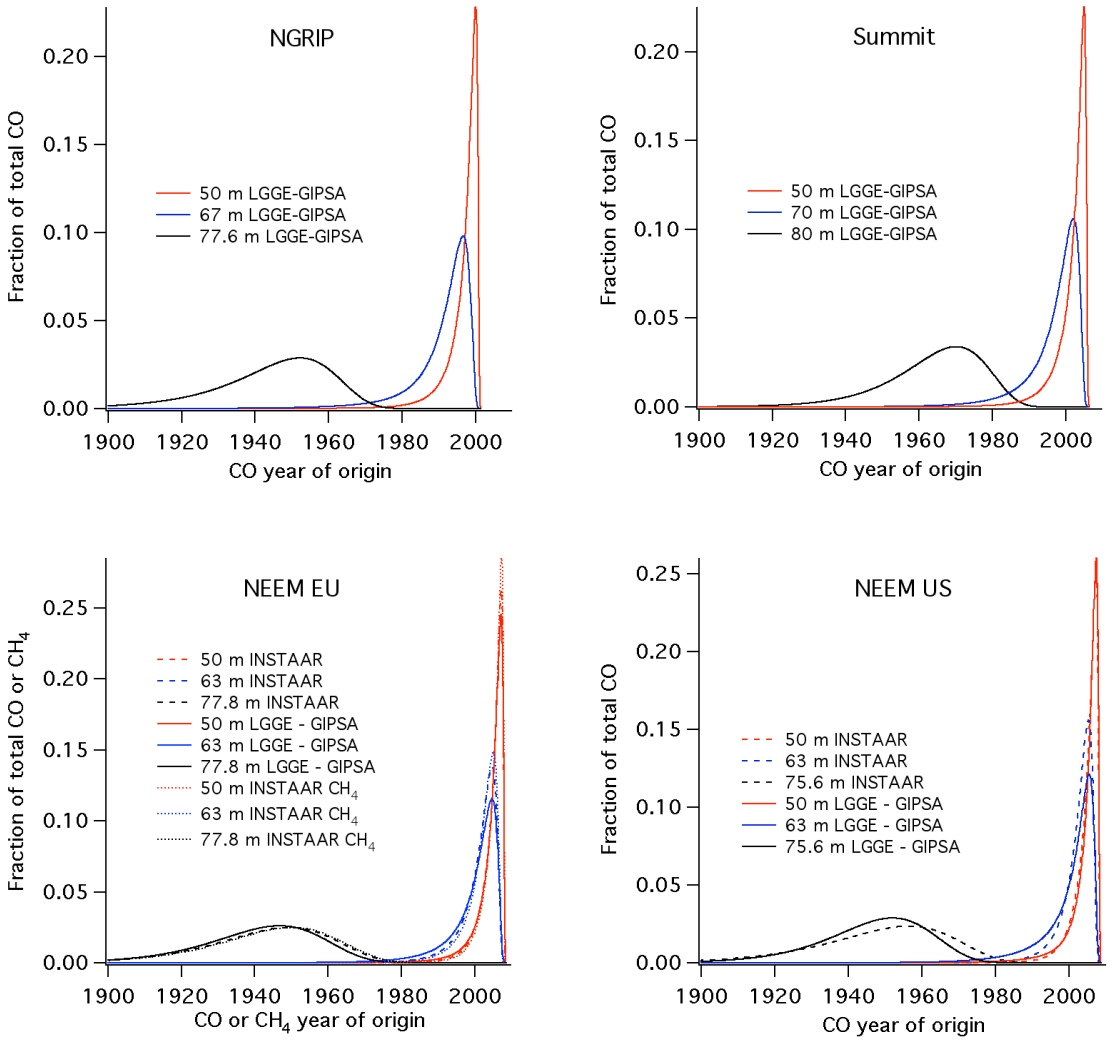

Figure S2. CO age distributions predicted by the LGGE-GIPSA and INSTAAR forward models for the different boreholes. For each borehole, curves are shown for one depth above the lock-in zone (red), one depth right near the start of the lock-in zone (blue), as well as the deepest sampled depth (black). The CH_4 age distributions predicted by the INSTAAR model are also shown for the example of NEEM EU borehole. As can be seen, the INSTAAR model CO age distributions are almost indistinguishable from the INSTAAR model CH_4 age distributions, demonstrating the validity of the CO-CH_4 comparison in Figure 6.

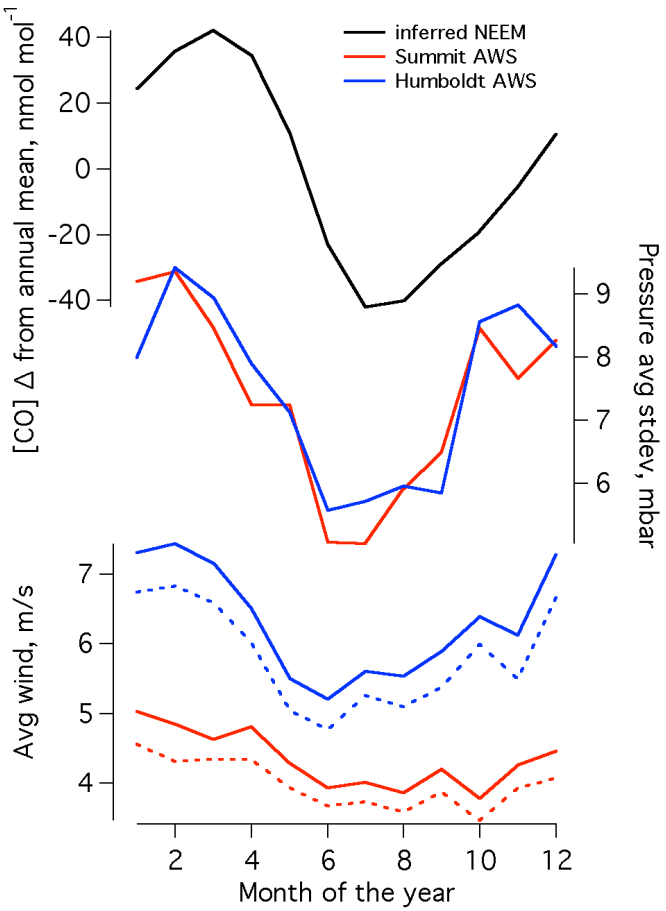

Figure S3. The seasonal variability of [CO], surface pressure and wind at Greenland sites. The monthly [CO] deviation from mean annual is estimated for NEEM, based on the existing NOAA Alert and Summit records up through 2008 (ftp://ftp.cmdl.noaa.gov/ccg). "Pressure avg stdev" is a measure of surface pressure variability and is calculated by finding the standard deviation of pressure values for each month, then averaging these standard deviation values over all Januarys, Februarys, etc. Each AWS has two wind speed gauges, and average wind speed from the 2nd gauge is plotted as a dotted line. Both Summit and Humboldt AWS records start in 1996. Humboldt wind data before 1996.35 and pressure data from May 1999 were excluded because of problems with gauges.

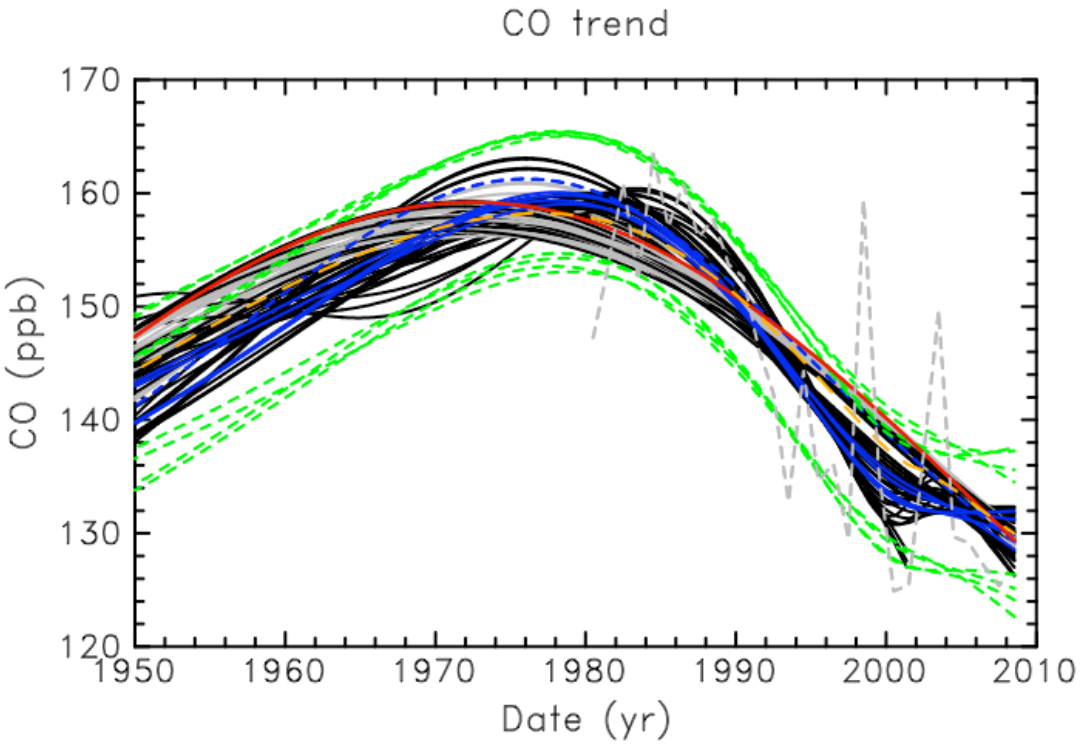

Figure S4: Comparison of the original set of 61 scenarios (black and grey continuous lines) with multi-site scenarios based on the new robustness-oriented optimal solution (in blue and red). Grey lines show the scenarios which would be excluded if the NGRIP biomass burning influenced data point at 42.6 m was taken into account. Red lines show the NEEM-EU + NEEM-US constrained scenario obtained with the INSTAAR model. Dashed blue lines show the NEEM-EU + NEEM-US constrained scenario obtained with the LGGE-GIPSA model. Continuous blue lines show the NEEM-EU + Summit + North GRIP and NEEM-US + Summit + North GRIP constrained scenarios obtained with the LGGE-GIPSA model. The dashed green lines show the model-determined uncertainty on the scenarios constrained with 3 sites (continuous blue lines). The long dashed orange line shows the average of the 61 selected scenarios in the manuscript (shown in black and grey here), it is the best estimate [CO] scenario used in Wang et al. (2012). The dashed grey line shows the annual mean [CO] atmospheric trend at Barrow (combined NOAA + OGI data, with OGI data adjusted to NOAA scale as explained in Section 3.3 of Supplementary Text).

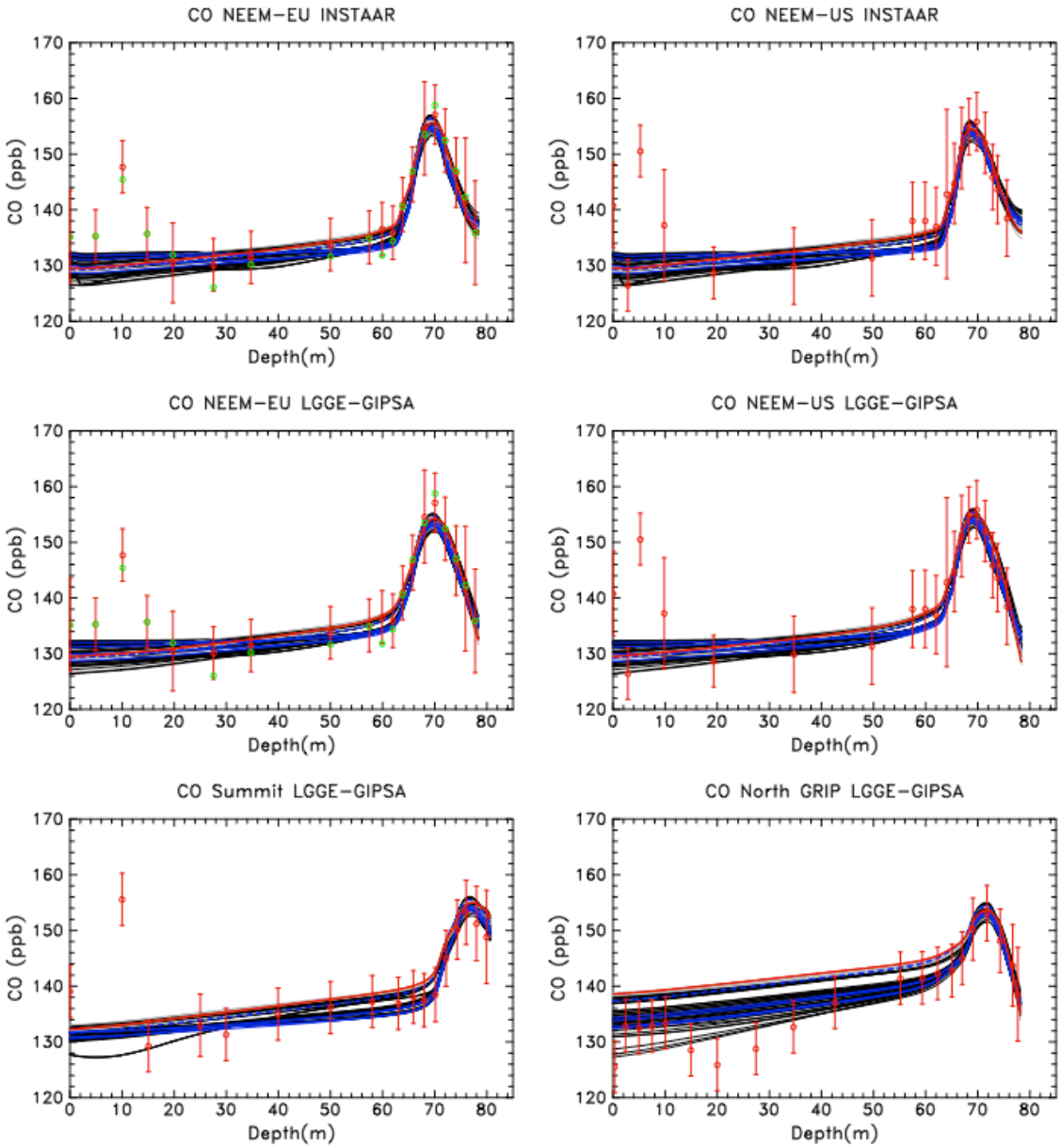

Figure S5: Convolutions of the successful [CO] scenarios with the transfer functions for individual forward models and boreholes. Finalized data sets are shown as red circles, with error bars showing the full uncertainties for the data-model comparisons as described in Section 4.2 of main text. Upper firn [CO] data are corrected for seasonality. Green circles show the NEEM-EU dataset modified to include Stony Brook data. Color-coding of the lines is the same as on Figure S4. Results using the original 61 successful [CO] scenarios of the main paper are shown in black and grey; the grey results would be excluded if the North GRIP data point at 42.6 m depth was taken into account. The blue and red lines show the results obtained from scenarios calculated with the new optimal solution of Witrant and Martinerie (2013).