Response to Referees – cp-2023-9

We are thankful to Murat Aydin and Vas Petrenko for their constructive comments on our article. We listed below our responses to the major and minor specific comments. Murat Aydin's comments are shown in blue, Vas Petrenko's in green. Our corresponding responses are below in black. For clarity we have regrouped different comments of the two referees, to provide a common answer. All the section and figure numbering included in our response to reviewers refers to the updated manuscript and SI.

Xavier Faïn on behalf of all co-author

General comments from Murat Aydin :

The manuscript titled "Southern Hemisphere atmospheric history of carbon monoxide over the late Holocene reconstructed from multiple Antarctic ice archives" presents new firn air and ice core measurements of CO from multiple Antarctic sites. Measurements from all sites are evaluated collectively to reconstruct a CO atmospheric history that covers the last three thousand years. All sampling and measurement methods, including the firn air modeling required to retrieve atmospheric gas records from firn air data, are sound and explained in sufficient detail here as well as in preceding publications. Both the firn air and ice core CO data and the retrieved atmospheric history are original and of interest to ice core community and atmospheric scientists.

I have to say I am not as enthusiastic about the interpretations offered in the manuscript as I am about the measurements. First off, the interpretations are overly qualitative. Given the heavy emphasis on the measurements, I am accepting of the fact that there is no atmospheric modeling, hence no real attempt to interpret the record quantitatively. However, it is strange that there is not even a quantitative overview of the present day CO budget. Section 3.5 and its subsections would benefit from a comprehensive reimagining.

General comments from Vas Petrenko

The authors present a compilation of Antarctic paleoatmospheric CO measurements from 7 firn air campaigns and three ice cores measured at high resolution using continuous flow analysis. To the best of my ability to tell, none of these data (with the exception of Berkner Island firn air) have previously been published. These new records span the time range from -835 CE to the 2000s and link with modern atmospheric observations in Antarctica. Two firn air models are used to reconstruct the CO history for the part of the record covered by firn air.

The records appear to be of very high quality, with careful attention given to calibration and corrections for analytical effects. The robustness of the reconstructions is confirmed by good

agreement among the records, as well as by agreement between firn air measurements and direct atmospheric observations for overlapping time intervals. The Antarctic ice core records also appear to be free of significant amounts of in situ – produced CO that has previously been seen in Greenland ice cores.

In my opinion this study represents a large advance in our understanding of atmospheric CO history. CO is a key player in global atmospheric chemistry, and a reliable CO history is required for a full understanding of natural variations in atmospheric chemistry as well as impacts of the industrial transition. This study provides this much-needed history.

The interpretation included in the paper is qualitative only, but I think this is OK as the reconstruction itself is a very significant contribution that merits publication in CP.

One of the most significant findings of the study is that prior Antarctic ice core CO measurements (presented in Wang et al., 2010) appear to have been biased high by 0 – 20 ppb (depending on ice core and time interval). I would trust the results of this new work over the older measurements due to improved measurement techniques, careful attention to procedural effects, agreement among multiple records as well as the fact that the continuous flow technique used here has been verified against at least one discrete-sample technique (Fain et al., 2022).

MAJOR COMMENTS

1. About the comparison with the Wang et al. (2010) study

Before the comments about section 3.5, I will list a few about the rest of the manuscript. The paper convinced me that the previous ice core CO record by Wang et al. (2010) includes spurious features, most notably the large biomass burning peak during the late 19th century probably never was. You speculate that analytical bias during discrete analysis could be the cause of the discrepancy. First, the Wang et al. (2010) data do not look biased because there is a period in the middle when the two data sets agree.

It is correct that a systematic bias of the Wang et al. [CO] dataset is not possible, because there is a period in ~1600 CE where both dataset agree. The manuscript has been corrected to remove all references to such bias. E.g., the following statement (line 573) was removed: "*We could speculate as a possible explanation for the difference that the quantification of CO contamination through the different discrete methods used for the D47/VST/SP ice cores significantly underestimated the true analytical bias introduced by these methods.*"

Second, Wang et al. (2010) also presented stable isotope data, with the higher CO mixing ratios corresponding to heavier isotopic ratios. This was interpreted as a combustion signal. Is there any information in the isotopic ratios that could be helpful in determining the source of spuriously high CO?

We are not aware of any interpretation of the isotopic CO signals that could help disentangle information about CO budget and possible lab or archives artifacts.

I think this new paper should go somewhat further in their discussion of the prior Wang et al results in the sense of alerting the readers that the isotopic measurements in Wang et al should now be interpreted with more caution – extra 10 – 20 ppb represents an additional 25 – 50% of CO, and this extraneous CO could certainly have a large impact on CO isotopic values. I still think that the Wang et al conclusion that the LIA CO minimum was mainly driven by reduced biomass burning is likely correct however.

We agree with this comment. Although we can not explain the [CO] patterns reported 13 years ago by Wang et al. (2010), we should alert the readers for a careful interpretation of these data. The following sentence was added to the manuscript (lines 573-575) :"*The SP/D47 CO isotopic ratios published by Wang et al. (2010) could also be impacted by additional, extraneous CO and thus should also be interpreted with caution.*". We also modified Sect 3.5.3 (line 742).

2. About Model descriptions

It is not clear how the mean and the uncertainty band for the firn histories presented in Fig. 1e,f are computed and then translated to Figs. S7 and S11?

The manuscript now provides more details about the mean and the uncertainty band for the firn histories presented in Fig. 1e,f, Figs. S7 and S11.

About the CSIRO model:

- In the main manuscript the sentence :

  "Uncertainties in the CSIRO firn reconstruction are calculated by a bootstrap method (Trudinger et al., 2016) that incorporates uncertainty in the firn measurements, the firn model parameters (using an ensemble of firn Green's functions), the Mawson atmospheric record and the ice reconstruction (when used)." (lines 242-246)

  was reworded as follows :

  "Uncertainties in the CSIRO firn reconstruction are calculated by a bootstrap method (Trudinger et al., 2016), whereby the inversion is repeated many times with firn measurements, the Mawson atmospheric record and the ice reconstruction all randomly perturbed according to their uncertainties, and used with firn Green's functions drawn randomly from an ensemble of Green's functions for each site (chosen during firn model calibration to represent uncertainty in the firn model parameters). The uncertainty

calculation is discussed in more detail in Supp Sec 2.4. The best-fit atmospheric history is calculated without any perturbations."

- The following sentence was added in SI Sect. 2.4.1, along with Fig. S11 :

"Figure S11 shows modelled concentration-depth profiles of CO in firn air from the CSIRO model. The solid lines correspond to the best fit cases in Fig 1f. The 2-sigma uncertainty band is generated by calculating firn depth profiles from the multiple atmospheric histories that result from the bootstrap method, thus corresponding to the uncertainty range in Fig 1f."

About the IGE-GIPSA model:

The uncertainty envelope of the IGE-GIPSA model combines a modeling error estimated as the size of the covariance matrix and a data prediction error reflecting the root-mean-square deviation between model results and firn data (Rommelaere et al., 1997; Witrant and Martinerie 2013). For single-site inversions, it generally reflects the variability in the data that the model cannot represent due to the smoothing effect of diffusion, and is consistent with an envelope of ± 2 sigma. In the case of multi-site inversions, it also reflects the model's difficulties in reconciling the different data sets.

The degree of smoothing in the IGE-GIPSA model is clearly higher than the CSIRO model. Are there differences in model physics in the lock-in zone, or are we mainly looking at the effects of different tuning approaches? You can address this by including firn model parameters in the SI, perhaps. The IGE-GIPSA model inversion displays a clear peak in the late 1980s followed by a minimum in early 2000s. I'm thinking this is primarily due to minor offsets in the data sets, is this correct? In the CSIRO inversions and the instrumental record (including Fig. S10) these features are not all that apparent.

The differences in smoothing in the inversion results is mainly due to different regularisation parameters and approaches taken in the inversion calculations, although there could be a small contribution from the different Green's functions of the observations used. An inversion of firn data needs some degree of regularisation to prevent unrealistic solutions with wildly oscillating atmospheric levels, as demonstrated by Rommelaere et al. (1997). Due to the different approaches to regularisation taken by each inverse method, there are not common regularisation parameters that can be compared. As indicated lines 194-196 of the manuscript, the method used in the IGE-GIPSA model emphasizes robustness (Lukas, 2008 ; cf. Sect. 2.4), and thus tends to lead to a smoother solution than other methods. Given that methodological choices in regularisation can lead to different degrees of smoothing of reconstructed atmospheric histories, it is an advantage in this study to be able to present two independent methods, and their differences reflect part of the overall uncertainty. Sect. 3.3.1 was modified to discuss the differences in smoothing in the inversion results.

Overall, it is important to note that the uncertainty envelopes of the IGE-GIPSA and CSIRO model results largely overlap, indicating that the difference between the two model results is not significant (Supplementary Figure S12). The different tuning approaches may play a role because the higher smoothness of the optimal solution from the IGE-GIPSA model implies that the slope changes of the solution are slower, and the existence of multiple mathematical solutions to the inverse problem renders the reconstruction of the amplitude of a peak difficult. Besides using different models and tunings, a major difference between the two approaches for that period is that the CSIRO modelling used the Mawson atmospheric data (starting in 1992) as a direct constraint whereas IGE-GIPSA did not. In addition, different firn datasets were used by the two groups. The CSIRO model was constrained exclusively with measurements made at CSIRO, whereas data from different laboratories were used to constrain the IGE-GIPSA model (see Sect. 2.3). Thus the IGE-GIPSA model results may be more subject to calibration biases but no systematic mismatch between CO data and model results at a single site is observed (Figures 1a and S7), calibration biases are thus likely small. Our main conclusion is thus the overall consistency, within uncertainties, of the results obtained with different datasets, models and tunings (Figure S12).

3. About the comparison with charcoal dataset and other proxies (Sect. 3.5)

Figure 5: Instead of just showing charcoal index, I think it would be more useful here to have a multi-panel figure showing the full new CO record together with acetylene and ethane (Nicewonger et al papers cited in this paper) in addition to charcoal.

We have modified Figure 5 and Sect. 3.5 to show ethane Antarctic record (Nicewonger et al., 2018) as well as charcoal records from intertropical and SH extra-tropicals areas. We discuss below these modifications in more detail.

It is not clear why you chose the two regional charcoal indices as the only benchmarks for comparison to the CO record in section 3.5.2. When calculating how much of the CO over Antarctica is from fires, the emission magnitudes from different regions cannot be ignored. Therefore, you need more than just sensitivity arguments to discard Africa and most of South America from consideration. It is worth thinking about what inference can be drawn from a comparison with the charcoal records presented in Fig. 5 within this context. CO lifetime is short but not short enough to just ignore the low latitude SH, or even all tropical fires. This is what model sensitivity studies show (e.g. Nicewonger et al., JGR, 2020). The most straightforward interpretation of the fact that neither one of the charcoal indices looks like the CO record during the LIA is that CO variability over Antarctica is driven by integrated fire variability occurring over a much larger geography. Most (80-90%, see GFED inventories) of the global fire emissions today occur at tropical latitudes (30degN – 30degS). I do not see any reason to think the situation should have been much different in the past because that is where most of the biomass is. For the same reasons, one can argue that those charcoal indices do not correlate

We agree with the reviewer that showing charcoal records from only South America and Australia regions was probably an incomplete picture. It showed, however, that [CO] in the Antarctic atmosphere was not only driven by one specific SH area.

We have completely reworded and merged Sections 3.5.2 and 3.5.3 to address the comments of both reviewers:

- Fig. 5 has been modified to include the ethane Antarctic dataset from Nicewonger et al. (2018), and the discussion emphasizes this comparison. CO and ethane have indeed very similar lifetime. We are not showing the acetylene record from Nicewonger et al. (2020) in Fig. 5 considering the shorter lifetime of acetylene (few weeks) compared to CO. However, the discussion still mentions this acetylene record.
- Fig. 5 has been modified to include charcoal composites from (i) the intertropical band (defined here at 25°N - 25°S), and (ii) the SH extratropical band (i.e., 25°S - 60°S). The SH extratropical band includes both South America and Australia, but also southern Africa.
- Section 3.5 now discusses the sensitivity of our [CO] atmospheric history to tropical fire emissions.
- The modified Fig. 5 is shown below :
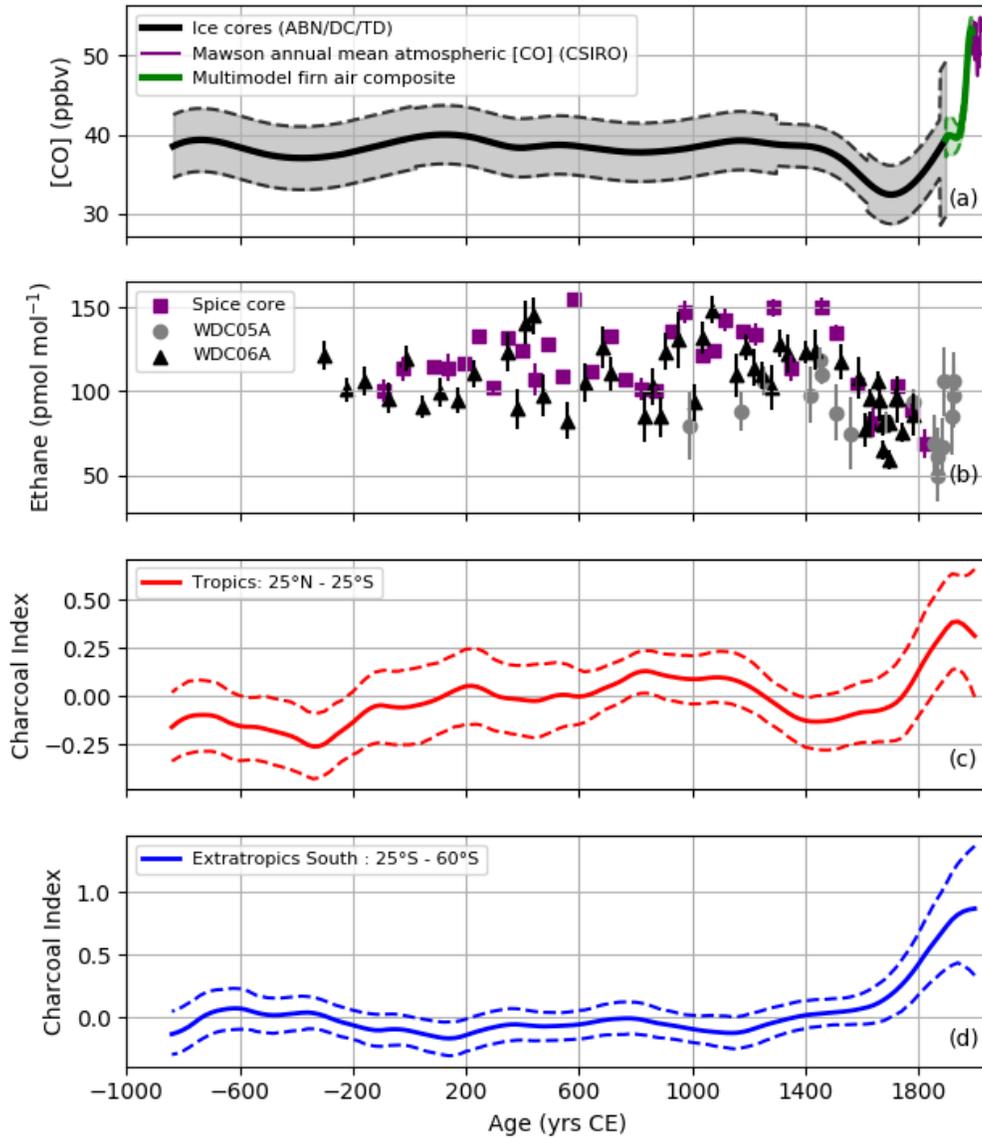
Figure 5. Antarctic Ice core and firn air CO (panel a; this study), and ethane (panel b; Nicewonger et al. , 2018) records. Charcoal indexes for the intertropical 25° N-25° S latitudinal band (panel c) and the extratropical SH 25° S – 60 °S (panel d) for the last 3000 years, extracted from the Global Paleofire Database (https://database.paleofire.org). Charcoal indexes are average Z-scores of transformed charcoal influx per region (100 yr smoothing window/1000 yr bootstrap). Dotted-line envelopes on the charcoal indexes represent the upper and lower 95% confidence intervals from the bootstrap analysis.

In contrast, the paper concludes the two charcoal indices may represent local fires instead of a regional signal (section 3.5.3, lns 669-673). I do not think that you can reach this conclusion using CO data alone for reasons that I stated above;

This statement was not based on CO data, but rather on the intrinsic nature of the charcoal data; it was likely incorrectly worded. However, it has been removed from Sect. 3.5 which does not discuss specifically South America and Australia charcoal records anymore.

however, let's assume for a moment this is true. What does this mean for the period prior to LIA? Can we draw any conclusions from the similarity between these charcoal records and the ice core fire proxies like CO, ethane, acetylene, and BC? I should note, I do not even agree that the charcoal indices do not show variability prior to the LIA. Most obviously, the Oceania index displays a long term rise that starts around 1000 CE and continues for about 1000 years.

Sect. 3.5 does not discuss South America and Australia charcoal records anymore.

I agree with Vas Petrenko that comparing your CO record to the other ice core fire proxies that are in agreement with CO would have put you on a better path to establishing, at the very least, that the Antarctic CO record prior to industrialization is indeed largely a record of gas emissions from fires. This is actually a pretty strong confirmation of the arguments you lay out in the beginning of this section about OH, methane oxidation, and BVOCs not factoring into CO variability in the PI atmosphere. In its current state, the manuscript is not very convincing in this respect. It is especially the comparison with ethane that would be of interest given that both gases have very similar lifetimes. Depending on what you see in those comparisons, I can see the discussion evolving in different directions, including some simple quantitative comparisons.

We have included the Antarctic ethane record from Nicewonger et al. (2018) in Fig. 5, and Sect. 3.5 now emphasizes the ethane - CO records comparison.

4. CO record during the industrial era

Finally, you mention in the abstract that both the growth from 1940 to 1985 and the following decline are observed globally (lns. 47-48), which is interesting except you do not offer any interpretation for the period after 1900 CE until the very last paragraph of the manuscript under summary and conclusions where this discussion seems misplaced. It would be better if there is a separate subsection about this period, discussing in a little more detail what you are referring to in the abstract.

We have added a subsection (3.6) (and a sentence in the abstract) to discuss the trend and likely causes in SH atmospheric [CO] during the industrial era. Specifically, our SH [CO] record

is compared with the recent NH [CO] reconstruction by Faïn et al. (2022) and previously Petrenko et al. (2013), as well as CO anthropogenic emission patterns (Hoesly et al., 2018) and inverse model inferred emissions since 2000 (Zheng et al., 2019). In this section, we mention the need to compare our atmospheric record with outputs from state of the art global chemistry-climate models. However, modelled dataset are not included in the manuscript, because we are preparing a third paper which will actually focus on such comparison, at a global scale (i.e., comparing both the NH [CO] record from Fain et al. (2022), and our SH [CO] reconstruction, with CMIP5 and CMIP6 model outputs).

MINOR COMMENTS

It is also worth commenting on the uncertainty in the ice core chronologies for D47 and SP.

This was an excellent comment. Wang et al. (210) report dating uncertainties of +/- 20 yrs (res. +/- 100 yrs) for their D47 (resp. SP) chronologies. D47 gas chronology was established by Barnola et al. (1995) using volcanic horizons, an ice flow model (the accumulation rate varies with depth at that site) and gas content measurements at close off depth. Dating the South Pole ice core is based on the Tambora volcanic horizon revealed by the ECM record and an ice age-gas age difference of 950 years determined by Schwander and Stauffer (1984).

The manuscript was modified (Sect. 3.4) to point out the large uncertainties on SP and D47 dating, which may explain the difference in chronology for the minimum [CO] levels during the LIA, between the Wang et al. (2010) dataset and our study.

What are the relative contributions of different sources of error to the uncertainty bands reported for the ice core data sets in Fig. 2 and what approach did you use in translating them to the uncertainty bands shown in Fig. 3 and Fig. 4?

Uncertainties are determined for each ice core record by considering three independent sources of errors : errors on solubility calibration factors, error on [CO] blank, and external precision of CO CFA measurements. This is stated in lines 441-442 of the manuscrit.

For some specific periods, more than one record is available. In that case, site specific uncertainties are combined following Tarantola (2005): assuming that we have two independent estimates with uncertainties of the same quantity x (e.g., x1±s1 and x2±s2), we can combine them and obtain the combined uncertainty S using the formula $1/S^2 = 1/s1^2 + 1/s2^2$.

The reference to Tarantola (2005) was added to the manuscript (Sect. 3.2.2).

Also, following a recommendation provided by H. Fischer during the preliminary stage of the editorial process, we now explicitly state in the caption of Figure 2 that the noise displayed by

the solid lines only represent the internal precision, while the envelopes report 2σ uncertainties combining uncertainties on CO blanks, solubility calibration factors, and external precision of CO CFA measurements derived from the reproducibility measurements.

Table 1 should include analysis dates.

Table 1 was modified to include analysis dates.

Line 103: Because Greenland ice core [CO] is so strongly affected by in situ production, absolute NH paleoatmospheric values are difficult to estimate. I think it would be more accurate to say that the Greenland ice core records allow for the reconstruction of "atmospheric trends" rather than "atmospheric history"

Line 103, we have replaced the wording "atmospheric history" with "atmospheric trend".

Table 1: should be "-44" for T at ABN

Table 1 was corrected as suggested.

Line 161: air stored in electropolished stainless steel tanks is affected by slow [CO] growth in my experience. If measurements from these tanks are used for [CO] reconstructions, it would be useful to see results of tests of [CO] stability in the tanks that were used.

The electropolished stainless steel tanks showed growth of typically 0.003 ppbv/day or less. The manuscript was modified accordingly.

Line 200: The Wang et al 2012 and Petrenko et al., 2013 references cited here are missing in the references list at the end of the manuscript

These two references were added to the list at the end of the manuscript. We also checked the list of references and we added a few others that were missing.

Line 340: correct this sentence for grammar

The sentence "Contamination resulting from entry of ambient air into the analytical system as breaks in the core was encountered." was corrected as follows : "The occasional entry of ambient air into the analytical system when core breaks are encountered could cause contamination."

Figure 1: middle panels (c and d) are too busy in my opinion – too many lines to be useful for the reader. I would recommend showing the GFs for just 2 – 3 depths per site – perhaps just the deepest sampled depth and another depth close to the lock-in depth.

We understand the point of the reviewer, but we decided to not change Fig. 1, panel c and d. We agree that panels c and d are busy, but we consider that the reader doesn't need to see every individual line. Our aim is to give the reader a sense of the significant overlap of green functions for many measurements, and the different widths.

Line 434: "filtered to remove lab air infiltrations"

The manuscript was corrected as suggested.

REFERENCES

Barnola, J. M., Anklin, M., Porcheron, J., Raynaud, D., Schwander, J. and Stauffer, B.: CO_2 evolution during the last millennium as recorded by Antarctic and Greenland ice, Tellus, 47B, 264–272, 1995.

Hoesly, R. M., Smith, S. J., Feng, L., Klimont, Z., Janssens-Maenhout, G., Pitkanen, T., Seibert, J. J., Vu, L., Andres, R. J., Bolt, R. M., Bond, T. C., Dawidowski, L., Kholod, N., Kurokawa, J., Li, M., Liu, L., Lu, Z., Moura, O'Rourke, P. R. and Zhang, Q.: Historical (1750–2014) anthropogenic emissions of reactive gases and aerosols from the Community Emissions Data System (CEDS), Geosci. Model Dev., 11(1), 369–408, doi:10.5194/gmd-11-369-2018, 2018.

Lukas, M. A.: Strong robust generalized cross-validation for choosing the regularization parameter, Inverse Probl., 24(3), 034006, doi:10.1088/0266-5611/24/3/034006, 2008.

Martinerie, P., Raynaud, D., Etheridge, D. M., Barnola, J. M. and Mazaudier, D.: Physical and climatic parameters which influence the air content in polar ice, Earth Planet. Sci. Lett., 112(1–4), 1–13, doi:10.1016/0012-821X(92)90002-D, 1992.

Nicewonger, M. R., Aydin, M., Prather, M. J. and Saltzman, E. S.: Large changes in biomass burning over the last millennium inferred from paleoatmospheric ethane in polar ice cores, Proc. Natl. Acad. Sci., 201807172, doi:10.1073/pnas.1807172115, 2018.

Nicewonger, M. R., Aydin, M., Prather, M. J. and Saltzman, E. S.: Reconstruction of Paleofire Emissions Over the Past Millennium From Measurements of Ice Core Acetylene, Geophys. Res. Lett., 47(3), doi:10.1029/2019GL085101, 2020.

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., 102, 30 069-30 083, 1997.

Schwander, J. and Stauffer, B.: Age difference between polar ice and the air trapped in its bubbles, Nature, 311(5981), 45–47, doi:10.1038/311045a0, 1984.

Tarantola, A.: Inverse problem theory and methods for model parameter estimation, Society for Industrial Mathematics, 2005.

Wang, Z., Chappellaz, J., Park, J. Y. and Mak, J.: Large variations in Southern Hemisphere biomass burning during the last 650 years, Science (80)., 330, 1663–166, doi:10.1126/science.1197257, 2010.

Witrant, E. and Martinerie, P.: Input Estimation from Sparse Measurements in LPV Systems and Isotopic Ratios in Polar Firns, in 2013 IFAC Joint Conference SSSC, pp. 654–659, 5th Symposium on System Structure and Control, Grenoble, France, 2013.

Zheng, B., Chevallier, F., Yin, Y., Ciais, P., Fortems-Cheiney, A., Deeter, M. N., Parker, R. J., Wang, Y., Worden, H. M. and Zhao, Y.: Global atmospheric carbon monoxide budget 2000–2017 inferred from multi-species atmospheric inversions, Earth Syst. Sci. Data, 11(3), 1411–1436, doi:10.5194/essd-11-1411-2019, 2019.