

We would like to thank the three anonymous referees for their careful reading of the manuscript and many thoughtful comments. We very much appreciate these constructive recommendations, and strongly feel that we were able to address all the points raised. We report a point-by-point response to the critiques. To aid the reader(s), we have continuously numbered the reviewers' comments.

REVIEWER #1 CONCERNS:

(1) Calibration of the system

(1.1) The authors calculate an average difference between measured and known values of the STDs (52±26 ppb) and apply this as a fixed blank correction, implicitly assuming that the standard reproducibility issue is a bias (blank) problem, which is constant within a range of variability. First, it appears to me that the 52 ppb number is weighed towards the STD1 results, by virtue of the fact that STD1 was analyzed 19 times vs. only 4 times for STD2. What is done here amounts to a one-point calibration near the more dilute 50 ppb std when even the baseline CO trace is in 100-300 ppb range (spikes measure up to 1 ppm).

We agree with the reviewer that the full loop data for standard gases are inadequate for the establishment of a precise calibration, with only 4 measurements reported for the 95.4 ppbv standard. We also agree that our two standard gases exhibit low CO mixing ratios compared to CO levels measured along the NEEM-2011-S1 core. Note that we were not expecting such elevated CO mixing ratios from the NEEM-2011-S1 core when planning the analytical campaign, and our standard gases were chosen to be similar to atmospheric concentrations. We corrected the manuscript (section 2.3.4) to specify that additional standard gases, and more full-loop measurements, will be required for better calibration of future continuous CO analysis.

However, we were recently (September 2013) able to conduct additional tests on the DRI-CFA system to evaluate its CO procedural blank. UHP N₂ was mixed with degassed DI water, and run through the full loop. This analysis was repeated three times (on different days), and revealed CO mixing ratios ranging between 40 and 50 ppbv. Note that direct measurement of the UHP N₂ on the OFCEAS analyzer shows CO mixing ratios below detection limit. Although these tests were carried out two years after the NEEM-2011-S1 analytical campaign (due to logistical issues: the CFA setup is located in Reno, NV USA, while the CO analyzer is located in France), they provide a direct estimation of the CO procedural blank. Although we cannot use it to calibrate our data, due to the 2 years lag between the NEEM-2011-S1 analysis and these full loop measurements involving UHP N₂, we have included this information in the manuscript because it supports the magnitude of our blank correction.

(1.2) What I see in the full-loop results shown in Table 1 can also be explained by a combination an initial dissolution loss on the order of 40-50%, followed by a contamination on the order of 75-80 ppb. I'm not suggesting this is exactly what is happening in their system (hard to conclude with just two points) but it is a possibility. There is a big difference between subtracting 52 ppb from raw data instead of assuming 50% of the signal was lost before 80 ppb was added to it. If this were the right method to use, most measurements may be

underestimates. Now, is 50% a reasonable number for loss to dissolution? I don't know. Dissolution loss is 7% for CH₄. What would we expect it to be for CO, assuming solubility equilibrium?

The reviewer describes above a direct interpretation of the 2-standards absolute calibration (i.e., blank of ~75ppb, losses of ~50%). However, a 50% loss of CO driven by solubility is physically unrealistic. A quick back-of-an-envelope calculation allows for estimation of the solubility of CO:

The solubility of CO, $L(\text{CO})$, in H₂O at 20deg is:

$$L(\text{CO}) = 1.918 \cdot 10^{-5} \left[\frac{\text{mol}(\text{CO})}{\text{mol}(\text{H}_2\text{O}) \times \text{bar}} \right] = 29.8 \left[\frac{\text{mg}(\text{CO})}{L(\text{H}_2\text{O}) \times \text{bar}} \right]$$

With the density of CO of 1.145 g/l the solubility per ml is:

$$L(\text{CO}) = 0.0260 \left[\frac{\text{mL}(\text{CO})}{\text{mL}(\text{H}_2\text{O}) \times \text{bar}} \right]$$

The partial pressure of CO in the gas sample is:

$$P(\text{CO}) = 1\text{bar} \times \left[\frac{150 \text{ ppbv}}{1 \cdot 10^9 \text{ ppbv}} \right] = 1.5 \cdot 10^{-7} \text{ bar}$$

The solubility of CO in water for the given partial pressure of CO is

$$L^*(\text{CO}) = 0.0260 \times 1.5 \cdot 10^{-7} = 3.9 \cdot 10^{-9} \left[\frac{\text{mL}(\text{CO})}{\text{mL}(\text{H}_2\text{O})} \right]$$

We have about 10% gas content in the ice, and the flow melting rate is about 18 ml/min. Thus dissolution happen from a ~1.8 ml/min gas volume to a ~16.2 ml/min water volume.

In the 16.2 ml water of the calibration mixture x ml of CO are dissolved:

$$x = 16.2 \times L^*(\text{CO}) = 6.3 \cdot 10^{-8} \text{ ml}$$

1.6 ml air of the calibration mixture initially contain a volume y of CO:

$$y = 1.6 \times \left[\frac{150 \text{ ppbv}}{1 \cdot 10^9 \text{ ppbv}} \right] = 2.4 \cdot 10^{-7} \text{ ml}$$

The dissolution for CO ratio of x/y is 26%. This result is already significantly below 50%. Furthermore, Rhodes et al. (2013) demonstrated that experimentally observed CH₄ dissolution losses are lower than when theoretically predicted. Applying similar calculations as above for CH₄ shows that CH₄ dissolution should reach ~23% (considering a 700 mixing ratio for CH₄ in the gas samples). Rhodes et al. (2013) observed only a 7% losses with a full loop calibration, and confirmed this finding by comparison with discrete analysis. Therefore, it is unlikely that CO dissolution losses are larger than 10%.

However, we recognize that our calibration data (Table 1) do not allow for correction of solubility effect. This is now clearly stated in the manuscript. We also clearly mention the implication of this for the CO record, i.e., underestimation of the reported concentrations.

(1.3) Any two points will yield a perfect linear line, no?

It is not clear for us where we use the wording “perfect linear line” in our manuscript. Unfortunately, only two standard gases NOAA-calibrated for CO mixing ratios were available for this study. We now recommend in our revised manuscript the use of at least three standard gases for future studies.

(1.4) If there were multiple calibration points, one can estimate the dissolution loss with some assumptions about where it is happening. As I mentioned above, they could easily be underestimating the CO levels, especially the high-end numbers (also says so on ln.4 pg.2840, contradicting these earlier statements).

As discussed in (1.2), we have evidence that our 2-point calibration slope based on gas standard measurements is not realistic. Furthermore, we discuss in (1.1) that the procedural CO blank is in the 40-50 ppbv range. Thus, we decided to not correct the data using a calibration line based on full loop results (Table 2). However, in doing so, we underestimate the CO mixing ratios (because we do not correct for solubility losses) as pointed out by the reviewer. This is now clearly stated in the manuscript, including in the abstract.

We believe that such lack of correction for solubility loss does not affect our analysis of the dataset for the following reasons:

- The replicate analysis described in section 2.3.3 demonstrates an excellent precision of our CO measurements. Thus, the entire dataset is remarkably internally consistent.
- The character of the data (abrupt spikes, elevated baseline) does not require an absolute calibration for discussion. All discussion reported in the manuscript is qualitative.

On ln.4 pg. 2840, we discuss CO mixing ratios in alpine samples, which are not related to the above.

(1.5) How do we define a reasonable slope, and if the slope is unreasonable, should we not explore why?

The calculation in (1.2) provides a back-on-the-envelope calculation, and demonstrates that 50% dissolution losses (slope) are not reasonable. The manuscript has been corrected to include this information.

(1.6) I would say the right thing to do is to add more points to the calibration curve and push its upper end to 500 ppb or higher (depending on what the outcome is for richer standards and ideally covering the entire measurement range).

We agree with the reviewer that more points would be required for the calibration. We have included this suggestion in the manuscript as a recommendation for future studies, but cannot improve the absolute calibration of the NEEM-2011-S1 core in retrospect.

(1.7) *If they don't want to make assumptions about when the loss and production takes places with respect to each other (i.e. which correction to make first), they can interpolate from the nearest standard runs to determine how much correction each measurement requires (better than extrapolating from 50 ppb).*

The CO measurements are affected by (i) solubility and (ii) contamination (i.e., elevated blank). A single standard measurement does not allow to distinguish between (i) and (ii). Furthermore, 80% of the full loop gas standard data available were obtained with the 50 ppbv standard. Thus, interpolation from a single and near standard was not an ideal solution.

(1.8) *There is verbiage in section 2.3.1 about regular flow of synthetic air-water mixture through the system in full-loop mode (the loop explanations come later but I presumed).*

In section 2.3.1, we explain that a mixture of in-house degassed deionized (DI) water mixed with a standard gas was provided to the sample line when we were no melting ice. This configuration avoids drying the system between melting runs : a mix of water and gas keep flowing through the Micromodule extraction device between melting runs. Rhodes et al. (2013) provides more details of this specific configuration, designated as the "tap loop". Full loop calibration, on the other hand, were limited by the amount of degassed DI water available in the bottle, and could not be run between each melting runs. The manuscript has been clarified.

(1.9) *Was this also done with a zero (no CO) tank to independently establish the blank? I cannot tell from the text. What were the blank levels observed in those experiments?*

See response to comment (1.1).

(1.10) *I acknowledge that the precision and the reproducibility of the system are good. I suspect this is why the authors chose not to address the calibration issue more thoroughly. I disagree with this logic. One can surely live with less than perfect accuracy, however, the issue is as much about the internal consistency as accuracy. Additionally, I think that one of the core arguments in the paper (the CO spikes being produced in situ) can only be proven with comparisons between well calibrated data sets (see my comments below). Calibration issues also hurt the discussions on the trends and comparisons with firn air and Eurocore measurements.*

We are aware that the absolute calibration of the continuous CO measurements from the DRI-CFA setup was not satisfactory during the NEEM-2011-S1 analytical campaign. This is the reason why we entitled the 2.3.4 section "an attempt for calibration". However, a qualitative discussion of the data is still possible, considering the good precision revealed by replicate measurements. We are working on improving this absolute calibration based on full loop, but unfortunately such improvement won't

apply to the NEEM-2011-S1 core. Future publications describing high resolution measurements of CO along ice cores will include a better characterization of contamination and solubility effects for CO.

All data were corrected for a constant offset of 48 ppbv (and we characterize the uncertainty on this offset to be +/-25 ppbv). This is a conservative approach, and we already reported in the manuscript that CO concentrations may be under-estimated for values well above the two standards (i.e., data are not corrected for solubility losses). However such approximation only affects absolute concentrations, and is thus not an issue to demonstrate in-situ production of CO considering the abrupt and elevated spikes observed (see also reply to comments (2)).

(2) In situ CO production

(2.1) I agree that there is strong evidence for in situ enhancement of CO in Greenland ice cores and all measurements may be elevated above atmospheric levels to a certain degree. However, can we single out in situ production as the only plausible explanation for the CO spikes with the given data, or even the trends in the baseline? Can we rule out the possibility that there is CO production during the melting and extraction from precursors that exist in the ice?

In order to address the reviewer comment, we have conducted additional measurements: we have replicated continuous CO analysis along a Greenland core collected at the D4 site, with different lab melting condition (light and dark melting). The CO record extracted from the D4 core is similar to the one we measured from NEEM-2011-S1. These results are now included as a new section (3.2.2) untitled “Does continuous laboratory melting drive CO production?”. No difference were observe when melting in the dark conditions, which suggest that photolytic production of CO does not happen during melting.

(2.2) I don't think replicate core reproducibility alone is sufficient for yes or no answers. The comparison of the baseline with the Eurocore data is suggestive of a common long term trend, I agree. But the truth is, there is no reason to leave out the spikes from this comparison (i.e. no reason to compare with a baseline line instead of a running mean) unless we are assuming the in situ mechanisms that lead to CO spikes are specific to the NEEM core. [...] Is there any independent evidence (ice chemistry) that can be used to show we should expect stronger in situ production of CO at the NEEM site than at Summit?

We agree with the reviewer that the only way to reconcile the NEEM-2011-S1 continuous CO record and the discrete Eurocore measurements by Haan et al., (1996) is to assume that in-situ mechanisms that lead to CO spikes differ between sites in Greenland.

Legrand, et al. (2013) have provided an up-to-date review of the existing data as well as new measurements of dissolved organic carbon (DOC) from a number of different sites using consistent methodology. These authors notably report that the lowest DOC concentrations at NEEM are higher than those at Summit (i.e., Eurocore).

We have used a single-case study to show that CO spike may be related to DOC enhancement in the NEEM ice. Although more data would be required to conclude, one plausible hypothesis would be that the NEEM ice is more DOC-enriched, and thus more prone to in-situ CO production.

(2.3) *Even without the spikes, CO measurements from NEEM are somewhat higher than the Eurocore ones (some caveat should be provided where ever they say the results from the two cores agree because I think they are only referring to general trends). This is true even during 1700-1800 when the effects of in situ production should be very low.*

The accuracy of our continuous CO measurements along the NEEM-2011-S1 core does not allow us to compare absolute baseline levels with the Eurocore record. This is now clarified in the revised manuscript. The overall conclusion is that the NEEM-2011-S1 CO record is not appropriate for atmospheric reconstruction, although it shows similarities with the Eurocore data.

(2.4) *It could be very useful to measure CO in discrete samples from NEEM (same analysis method as Eurocore). A limited number of discrete samples should be enough to see if discrete samples from NEEM would average over the continuous flow measurements like they should. I presume the capability exists given the measurements in Alpine ice cores.*

Unfortunately, the capability of measuring CO in discrete sample was not available during this study. Note that the alpine samples have been analyzed few years before the continuous NEEM-2011-S1 record. We have included the need of discrete measurements as an important improvement which would benefit future studies.

(3) Drilling fluid contamination

(3.1) *I'm not sure how close to the rind these ice core sticks were but it would be reasonable to expect some level of drill fluid contamination on outer layers of fluid- drilled cores. There may also be memory effects. I suspect the water solubility of the drill-fluid is fairly low and it could take a long flush to clean it all out. In a continuous analysis, this may lead to artifactual trends after a contamination event.*

One section of the core (249-292 m depth) was heavily cracked, and we cannot rule out that CO observed over this section is not impacted by drilling fluid contamination. Continuous chemical analyses show disturbed baselines over the 249-292 m depth range that could be attributed to drilling fluid.

However, both the main ice core stick and replicate stick (cuts M and B, respectively) were cut at a distance of at least 1cm of the core edge and as a result we do not expect drilling fluid contamination in uncracked ice sticks. Our DOC analyses support this expectation - DOC shows very low concentrations on uncracked samples, typically lower than 20 ppbC.

Overall section 3.2.1 has been clarified and rewritten. Grey shading has been added to Figure 3 to highlight the depth range potentially affected by drilling fluid contamination.

(4) Chemical processes that drive in situ CO production

(4.1) This is a difficult section to follow, especially the last three paragraphs (pgs. 2836--- 2837). There are many issues being addressed all at once and it does not work for me.

We have clarified the manuscript, by creating two new sections that address more specifically the different issues discussed.

(4.1) The firn or ice question (essentially where the enhancement occurs) and related discussions are scattered and confusing. If this is a firn process near lock-in, why would the firn data make sense? Is this not a contradiction with conclusions of earlier firn works by Petrenko et al. (2012) and Wang et al. (2012)?

This section does not discuss firn vs ice processes. Comparison of the NEEM-2011-S1 data with published firn air data (Petrenko et al., 2013) is now included in section 3.3.

(4.2) Is it possible to quantify the arguments for continual production based on the depth dependence of spike amplitudes, may be with a separate figure of deviation from the median vs. depth? I can see that they are more intense between years 400-800 but they get smaller again on the deepest layers.

We agree that spike amplitudes and frequency are interesting parameters to link to depth when investigating the hypothesis of continual production. We did observed an increase in spike occurrence with depth, and we report now this information in section 3.1.2 as follows : “occurrence of abrupt CO spikes increased with depth along the NEEM-2011-S1 core, with 29 and 38 spikes observed in the 80-200m and the 300-400m depth ranges, respectively.”

(4.3) Figure 4 needs tick labels for the upper x-axis since tick marks do not line up and the axis breaks are at a different positions.

Upper x-axis on Figure 4 has been corrected.

(4.4) I cannot really follow the discussions about which peaks line up and which do not. The figure also needs a better caption: which trace belongs to which axis etc.

We agree that this discussion was difficult to follow because the submitted figure 4 was not correct. We have corrected figure 4, which shows now CO in red, NH₄⁺ in blue, and rBC in black. It is clear that two CO-enhanced events are reported, with the only the second event showing concomitant increases in CO, NH₄⁺ and rBC.

(4.5) *Figure 5 axes should be reversed, CO is the dependent variable here.*

We understand that CO is the dependent variable, and thus would be better shown on y-axis. However, figure 5 has two plots distributed vertically (rBC vs CO, and NH_4^+ vs CO, respectively). We find figure 5 easier to read the way it is as both plots can share a common “CO” x-axis.

(4.6) *The answer (based on figure 5) is it is not clear (no statistically significant correlation?).*

Figure 5 allows us to investigate how CO concentrations relate to NH_4^+ and rBC levels in ice layers affected by deposition of pyruvic aerosols. Interestingly we observed a lack of dependency between these variables. This is an important result to report, although the reasons involved are still unclear. The manuscript text has been adjusted to include the lack of significant correlation between these variables.

(4.7) *After that, things gets nearly incomprehensible. Figure 6 and all the discussions centered around it do not provide conclusive evidence on what is going on, which makes me question whether figure 6 should be shown in the paper at all. Does it make sense to focus on one peak unless what we learn from it will lead us to the promised land?*

We hope that reorganization of the manuscript, including dividing the text in new sections, will make it more comprehensible. We do not pretend in our manuscript that investigating DOC and its speciation for two ice sections showing low and high CO, respectively, will explain all processes involve in the in-situ production of CO in the NEEM ice.

We propose a methodology to investigate the link between CO and organic content of the ice to explore the question of CO in-situ production. We report only preliminary results, and more studies will be required to resolve this issue. Unfortunately, the scope of our discrete DOC analysis was limited by the amount of ice available.

(4.8) *I would argue the real advantage of the higher resolution measurements is that it provides many instances of abrupt changes and increases the chances of finding a statistically significant correlation between the spikes (all of them) and one of the biomass burning markers*

The approach described by the reviewer here is exactly the one presented in figure 5. We considered all abrupt CO spike events along the NEEM-2011-S1 record, and investigated their correlation to available biomass burning markers (i.e., rBC and NH_4^+) No significant correlation was found, as discussed in the new section 3.4.1 and in response to comment (3.7).

(4.9) *Another thing that puzzles me is why there is hardly any focus on changes in baseline levels in this section? Would long term trends in pyrogenic aerosols not influence the long term trends in CO?*

This is a good point. We have investigated if the monotonic increase in CO with depth along the NEEM-2011-S1 core could be related to any similar trends in available biomass burning proxies. However, none NH_4^+ and rBC showed monotonic increasing trend with depth, similarly to CO. We are now reporting this result in section 3.4.1 of the revised manuscript.

(4.10) *There is a misplaced reference to figure 5 on ln.26 pg. 2836.*

The manuscript has been corrected.

REVIEWER #2 CONCERNS:

(5.1) *The investigators would likely benefit from diagnostic tests using CO-free synthetic air.*

See response to comment (1.1).

(5.2) *based on the spread shown in Table 2, it appears some of the contamination is due to a leak of room air into the system*

Table 2 present CO concentration measured in discrete alpine samples following the procedure of Haan et al. (1996). The large variability in CO concentrations reflects the ice content. Note that lab air is one order of magnitude lower in CO concentration than data reported in Table 2 for alpine ice, and thus cannot explain such CO elevated levels.

(5.3) *I don't know that I agree that spikes of elevated CO serve as a useful tracer of biomass burning, or any other geophysical process.*

This is also a concern raised by Reviewer 1 in their introductory comment. We accept that considering that in-situ produced CO as a potential proxy for biomass burning is too speculative at that point because we could not establish strong relationship between the CO record and other biomass burning proxies records along the NEEM-2011-S1 ice core. Thus, we removed all statements describing in-situ produced CO as a potential proxy for biomass burning history from the manuscript.

(5.4) *Measurement of the stable isotopes of CO might provide insight into the sources of elevated CO.*

This is an interesting suggestion. However, this would require (i) identification of the processes leading to CO production in the ice, and (ii) understanding of how such processes impact the isotopic signatures of CO. Finally, the available method to measure CO isotopic ratio in ice samples (e.g., Wang and Mak, 2010) require a large amount of gas. We observed abrupt CO spikes likely limited to

ice layers of only few cm. Such high depth resolution required for CO isotope analysis cannot be resolved with conventional ice core drilling.

REVIEWER #3 CONCERNS:

(6.1) *I recommend stating clearly, in the abstract or the conclusion (or both) that at this stage the full system is not suitable for meaningful past atmospheric CO reconstructions from ice cores [...]. The system would, however, be suitable for identifying ice layers that are promising targets for more accurate discrete measurements of CO and possibly CO isotopes.*

We have adjusted the manuscript to clearly state the limitations of the system and to suggest that that improvement of the blank characterization and correction is required before analysis of Antarctic ice.

(6.2) *While the presented evidence is strong that the measured CO record (specifically, intensity of in-situ CO) is imprinted by the intensity of biomass burning (particularly at Northern high latitudes), the suggestion that this could be used as an integrative proxy for past biomass burning is, in my opinion, over-reaching and should be removed.*

See response to comment (5.3).

(6.3) *The procedural CO blanks for the overall system do not seem well characterized.*

See response to comment (1.1).

(6.4) *Dissolution should be a very minor effect for CO, whose solubility is very close to that of air. A quantitative estimate of the dissolution effect could be obtained, for example, by taking the well-characterized CH₄ dissolution effect (as in Rhodes et al 2013) and considering the CO vs CH₄ solubility in pure water.*

See response to comment (1.2).

(6.5) *From Table 1, it seems to me that the system blank depends on the CO concentration of the air being analyzed – higher blank for lower CO. This suggests not only CO production (or desorption), but also CO consumption (or adsorption) in the system. A more complete characterization would involve more CO standards, with concentrations ranging up to the high values found in most ice core samples; this should be done if possible.*

See response to comment (1.1).

(6.6) *As far as I could understand, neither the “internal loop” nor the “full loop” procedural tests are fully representative of the overall procedural blank, because they include some components that are not in the real sample flow path, and exclude others (such as the melt-plate).*

We fully agree with the point raised by the reviewer. We have clarified this issue in the manuscript (section 2.3.4) by highlighting the need for discrete measurements to validate any CFA-based calibration.

(6.7) *While it is difficult to re-create a truly representative procedural blank test with this type of continuous measurement, could a more realistic assessment of the blank be obtained by measuring replicate ice samples on a discrete-sample system that has better-characterized blanks (such as the old Haan et al system or the newer Wang et al system)?*

See response to comments (6.6) and (2.4). Yes, discrete measurements will be mandatory to confirm any CFA-based CO calibration. Unfortunately, such discrete analysis capability was not available for this study. However, the revised manuscript emphasizes that combining both continuous and discrete analyses will greatly improve future investigations of CO ice core records.

(6.8) *While it does seem that there is evidence for drilling fluid affecting CO in at least some of the ice, the method of establishing this seems indirect and leaves much ambiguity in interpretation. Instead, could the authors try taking two replicate halves of a dry-drilled ice core section, dipping one of the halves in the drilling fluid, then cleaning and drying it (using regular protocol), and then proceed with cutting and analysis as usual? It seems that this kind of test would give a less ambiguous result. As it stands now, the drilling fluid question adds much uncertainty into the discussion of in-situ CO production.*

Although the approach suggested by the reviewer to evaluate the impact of drilling fluid on CO measurements is interesting, we do not consider that it is required. See response to comment (3.1). We have implemented different approaches to evaluate if/where drilling fluid contamination occurred, including discrete DOC analysis, comparison with continuous chemistry data, and core inspection for fractures and cracks. We consider that we have been able to clearly identify which sections of the NEEM-2011-S1 core have been contaminated by drilling fluid.

(6.9) *Wording and minor comments:*

- Abstract, line 12. Poor grammar – re-word this sentence

Completed.

- p.2824, line 23. “applying a depression” – do you mean a pressure gradient?

Yes, the manuscript was corrected.

- p.2826, line 19. “more than ten” – it would be more precise to state exactly how many

33 meters of replicate ice were melted. The manuscript was corrected.

- p. 2831, line 7. Is “FWHM” already defined in the paper?

FWHM means Full Width at Half Maximum. The manuscript now states this.

- p. 2831, lines 12-13. I don’t understand how two different depth ranges can correspond to the same time interval.

This typo was corrected.

- Fig 3 caption. For clarity, it would be helpful to state here that all the data are corrected for the estimated (and large) procedural blank

Completed.

- Fig 4 and caption. Please add a description of which line represents which tracer. Why not plot more of the data (CO vs BC and NH4)?

Completed.

- p.2831, line 20 – 23. Not that air in closed porosity is older than air in open porosity at the same depth level, so should probably contain lower CO than firn air measurements

This is true. However, (i) the CO patterns driven by in-situ production and (ii) the lack for an accurate absolute calibration did not allow us to investigate in details such difference between the continuous NEEM-2011-S1 CO record and the firn data published by Petrenko and others (2013).

References

Haan, D., Martinerie, P. and Raynaud, D.: Ice core data of atmospheric carbon monoxide over Antarctica and Greenland during the last 200 years, *Geophysical Research Letters*, 23(17), 2235–2238, 1996.

Legrand, M., Preunkert, S., Guillhermet, J., Faïn, X., Alekhina, I. and Petit, J. R.: Water-soluble organic carbon in snow and ice deposited at Alpine, Greenland, and Antarctic sites : a critical review of available data, *Climate of the Past*, 9, 2195–2211, doi:10.5194/cp-9-2195-2013, 2013.

Petrenko, V. V., Martinerie, P., Novelli, P., Etheridge, D. M., Levin, I., Wang, Z., Blunier, T., Chappellaz, J., Kaiser, J., Lang, P., Steele, L. P., et al.: A 60 yr record of atmospheric carbon monoxide reconstructed from Greenland firn air, *Atmospheric Chemistry and Physics*, 13(15), 7567–7585, doi:10.5194/acp-13-7567-2013, 2013.

Rhodes, R. H., Faïn, X., Stowasser, C., Blunier, T., Chappellaz, J., McConnell, J. R., Romanini, D., Mitchell, L. E. and Brook, E. J.: Continuous methane measurements from a late Holocene Greenland ice core: Atmospheric and in-situ signals, *Earth and Planetary Science Letters*, 368, 9–19, doi:10.1016/j.epsl.2013.02.034, 2013.

Wang, Z. and Mak, J. E.: A new CF-IRMS system for the quantification of the stable isotopes of carbon monoxide from, *Atmospheric Measurement Techniques*, 3, 1307, 2010.