General comments

In this study the authors present a synthesis of air pollution data from four different High Alpine European ice cores for the last ~250 years. They analyse how representative the single core data are for the common pollution history of the area. They create composite records of 10-year averages for the respective pollution species and calculate "times of emergence" (TOE) when the actual anthropogenic pollution becomes significant. This is an interesting topic and especially the investigation of the representativity of one ice core record for the atmospheric signal of an entire region is very important and was not sufficiently discussed so far. Also the findings of the different times of emergence of the pollution signals in Europe and the present day conditions, including partly failed abatement measurements, are highly relevant and should be published.

The paper is overall well written. The paragraphs follow a clear structure and are easy to follow. However, the general focus of the study remains a little unclear to me. The title and the statements at the end of the discussion (section 3) paragraphs rather suggest that the investigations of the TOE and the evaluation of the present day situation were the main objectives. On the other hand, the authors also dedicate a good deal of the study to the evaluation of the representativity of a single core record for the whole region. In my opinion, both discussions lack depth and critical evaluation. The paper would definitely benefit from a clarification of the main objective and a more focused approach (methodological or interpretation of pollution history). Especially the discussion section (3.2-3.4) is sometimes rather vague and needs some revision. I think the discrepancies between the single core records (Fig. 2) and the resulting consequences for the composite record (and the following interpretation) are not discussed and assessed sufficiently (see specific comments).

The main object of the study is to investigate, how representative one Alpine ice core is to document the pollution history of Western Europe. As suggested by the referee we will clarify this throughout the manuscript, where not done already (end of introduction, first chapter of section 3., conclusion). We think that the chosen title does not contradict the main object of the study and prefer to keep it. In chapter 3 we will clarify that for investigating the representativeness a) absolute concentration levels and b) long-term concentration trends of BC, Cd, F$^-$, NH$_4^+$, NO$_3^-$, Pb, and exSO$_4^{2-}$ from the four sites (CDD, CG, FH, and GG) were compared. For a more quantitative and less "vague" comparison of the long-term trends we added a correlation analyses (Fig. R1).
According to the comments of the referee we will rearrange the structure of the Results and Discussion part:
- 3. Results and Discussion
general paragraph:
*"Average emission sensitivities during the period 2000-2009 derived with FLEXPART, indicate that in general the four ice-core sites CDD, CG, FH, and GG are most sensitive to emissions from the countries surrounding the Alps, i.e. Switzerland, France, Italy, Germany, Austria, Slovenia, and Spain, without substantial discrepancies between the four sites (Figure 1). A noticeable extension of the pollution source area into Spain is recognized for the westernmost CDD, and into larger parts of*

*Germany for the northernmost FH site.*

*To investigate, how representative ice-core records from the four sites are to document the pollution history in this deduced source area we systematically compared a) absolute concentration levels (section 3.1.) and b) long-term concentration trends of BC, Cd, F⁻, NH₄⁺, NO₃⁻, Pb, and exSO₄²⁻ (section 3.2.).*"

- 3.1. Absolute concentrations
  (former chapter 3.1.2)
- 3.2. Trends
general paragraph:

*"Generally, an excellent agreement in the ice-core long-term concentration trends of all investigated pollutants is observed for CDD, CG, FH, and GG (Figure 2). This is confirmed by correlation analyses, revealing high correlation coefficients ($0.86 < r < 0.99$) between different records of the same compound (new figure R1). All correlations are significant at the 0.001 level.*

*The observed, characteristic differences between the concentration trends of the individual seven pollutants reflect dissimilarities in respective emission sources, emission trends, atmospheric chemistry of precursor species, and compound specific implementation and efficiency of emission abatement measures, which is discussed in detail in the sections 3.2.1-3.2.3. For individual pollutants this additionally includes discussion of common features and differences in the 10-year concentration records of the four sites and the consequences regarding the formation of Alpine composites and representativeness to capture the Western European pollution history. The latter is evaluated by TOE and comparison with available model estimates to precisely time the onset of increased pollution levels and model validation of air pollution trends, respectively."*

- chapters 3.2.1 – 3.2.3.

former chapters 3.2.-3.4., +discussion about consequences of common/different features for the formation of Alpine composites and representativeness will be added also considering the results of the correlation analyses
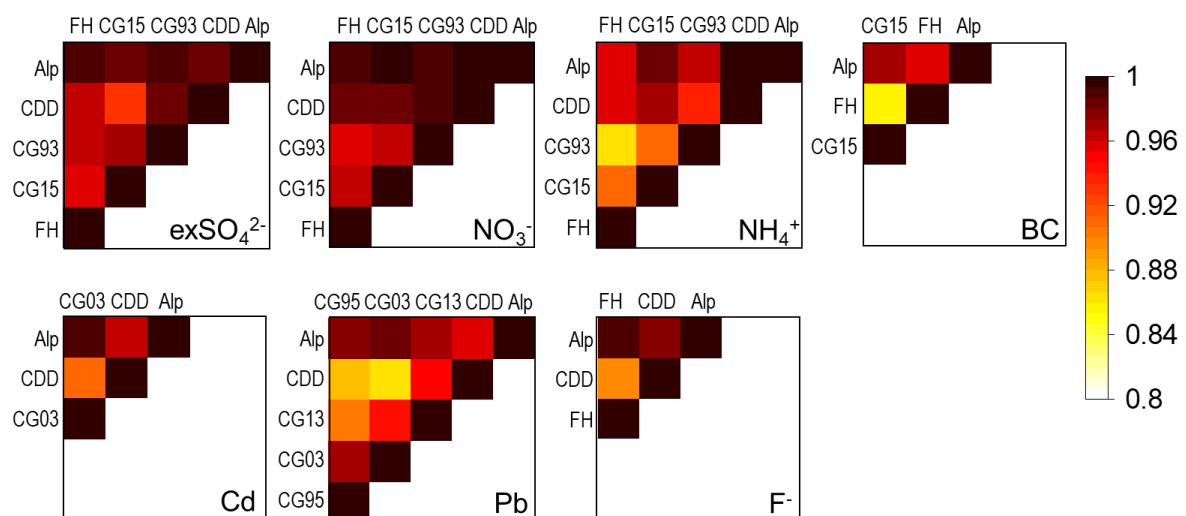


Fig. R1: Correlation matrices showing the correlation coefficients r between 10-year concentration records from different ice cores and the Alpine composite (Alp) for the individual pollutants exSO₄²⁻, NO₃⁻, NH₄⁺, BC, Cd, Pb, and F⁻

My other major point of criticism is the methodological approach for averaging in section 2.3. It is not justified why the authors calculate 10-year averages for the single ice core records. Is that indeed necessary to "average out inter-annual-fluctuations"? (L203). If only looking at a 250 year time range this seems to smooth out a lot of variation. Could another averaging window also be sufficient and maybe provide more detailed insights? This needs some statistical evaluation. The reasoning behind the 10-year averaging seems random, remains unclear so far and will have consequences for the following discussion of TOE and pollution records. This section needs to be extended, clarified and methodologically justified.

We agree with the referee that this point of the chosen averaging window was not explained properly and we will extend the discussion in section 2.3 and add a new supplementary figure R2. There are two main reasons for the chosen 10-year window: (1) averaging out inter-annual fluctuations mainly related to variations in pollutants vertical transport to the high altitudes controlled by atmospheric stability and (2) accounting for the low data resolution and high dating uncertainty in the lower ice core parts (e.g. 10-25 years for the FH in the period 1750-1800).
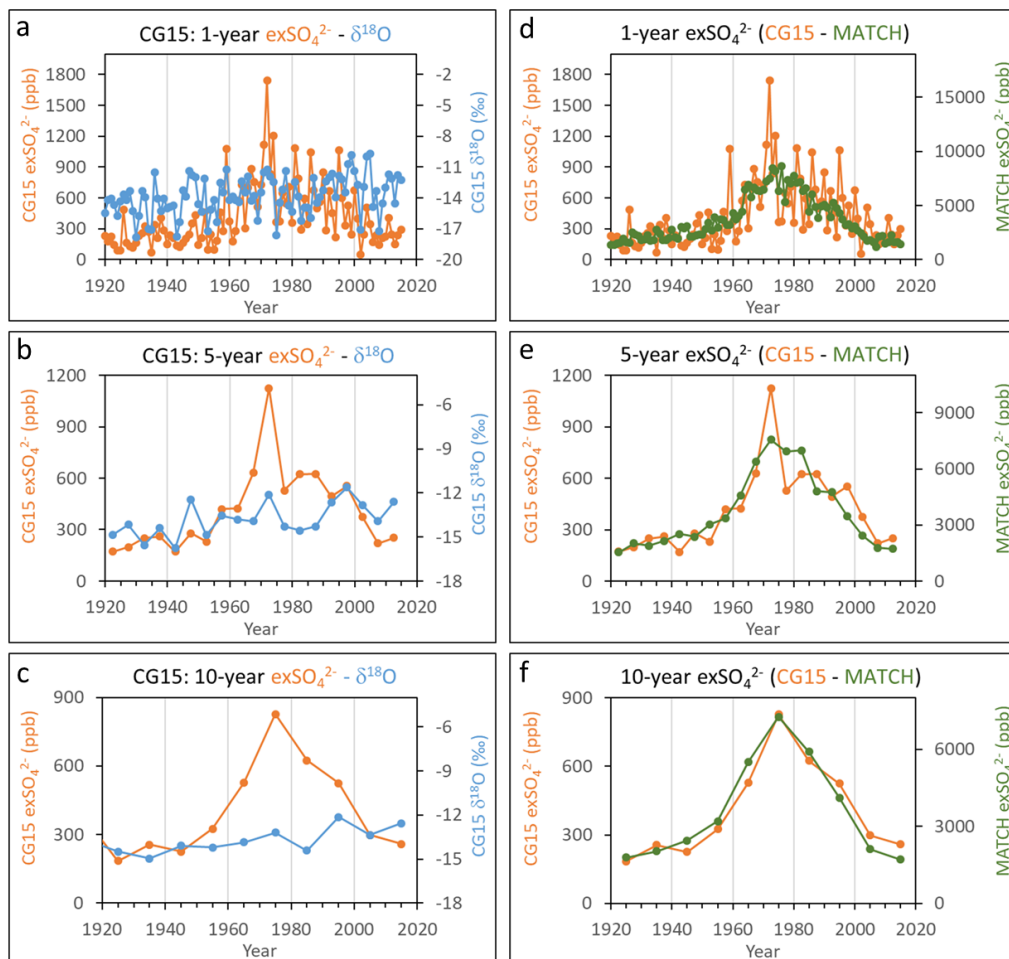


Fig. R2: Comparison of CG15 $exSO_4^{2-}$ concentration records with CG15 $\delta^{18}O$ records (panels a-c) and with modelled $exSO_4^{2-}$ concentrations (MATCH-ECLAIRE) (panels d-f) in the period 1920-2015 using different averaging windows (a,d – 1-year, b,e – 5 year, c,f – 10 year).

To select an adequate averaging period, we investigated the influence of the different averaging windows on the CG15 $exSO_4^{2-}$ record in the period 1920-2015 (Fig. R2). Figure R2a illustrates that

annual exSO$_4^{2-}$ maxima/minima are in part synchronous with $\delta^{18}$O maxima/minima. Since $\delta^{18}$O is a temperature proxy (Bohleber et al., 2013), this reflects a stronger/weaker vertical transport of pollutants from their source regions at lower elevations to the high-altitudes sites during warmer/colder summers. This is still partially visible in the 5-year means (Fig. R2b), e.g. the temperature ($\delta^{18}$O) maximum in the period 1970-75 lead to an amplification of the exSO$_4^{2-}$ concentration at CG compared to what is expected based on the SO$_2$ emission history (Fig. R2e). This modulation of the ice-core record by year-to-year variability of vertical transport/atmospheric stability also produces a larger variability compared to the modelled concentrations for the annual and 5-year means (Fig. R2 d,e), but not for the 10-year averages. The MATCH-ECLAIRE model does not have sufficient resolution to resolve vertical transport.

Thus, we argue that the ice-core 10-year averages (1) are representative for changing pollutant emissions, averaging out the inter-annual (short-term) fluctuations related to temperature dependent pollutant vertical transport to the high-alpine sites and (2) take into consideration the low data resolution and higher dating uncertainty in the pre-industrial period.

I have some specific comments, especially about some figures and the respective discussions.

Specific comments

L21: please put "long term" in perspective. For most ice core studies, 250 years would not be considered long term

We put the wording "long-term" already in perspective by adding the time period (AD 1750-2015) and think this is sufficient information for the abstract of the paper.

L21: please specify "several" ice cores

Will be specified: *"Two to five ice cores from four high-Alpine sites located in the European Alps analysed by different laboratories were used, depending on the data availability of the different air pollutants."*

L37f: This statement lacks evaluation. What are the consequences from only four core records corresponding well with the modelled trends? Please elaborate

We will add potential reasons: *"Only four ice-core composite records (BC, F$^-$, Pb, exSO$_4^{2-}$) of the seven investigated pollutants correspond well with modelled trends, suggesting uncertainties of the emission estimates or in the representation of chemical reactions in the model for the other pollutants."*

L51: change "ice cores" to "ice core records"

We will change this accordingly.

L66ff: Maybe a problem of wording: "… indicates common source regions for the different Arctic sites…" I think this is a rather generalized and bold statement, especially because in the next sentence you state that this is not true for e.g. Svalbard. Please clarify.

Will be changed to *"This finding indicates common source regions for the different Greenland sites in North America and Europe…"*

L79f: Reference for huge Himalayan dust deposition?

We will add a reference (Sierra Hernandez et al., 2018) accordingly.

L87: you could add: Preunkert, S., McConnell, J. R., Hoffmann, H., Legrand, M., Wilson, A. I., Eckhardt, S., et al. (2019). Lead and antimony in basal ice from Col du Dome (French Alps) dated with radiocarbon: A record of pollution during antiquity. Geophysical Research Letters, 46, 4953– 4961. https://doi.org/10.1029/2019GL082641

We will add this reference.

L90f: The sentence about the downward trend is somewhat repetitive throughout the manuscript. Please shorten.

Will be shortened and repetitions avoided.

L93: Please specify that "annual" concentrations at Colle Gnifetti are also heavily biased towards the summer season because of major wind erosion of winter snow.

This point is already detailed in the chapter 2.1.1.

L102: which were the "different laboratories"?

This is different for specific pollutants and is already mentioned in Chapter 2.2.

L114: Can you add a reference for the atmospheric lifetime of the species? Are they the same for all different pollutants?

Atmospheric lifetimes vary depending on the conditions (occurrence of precipitation). We think for the purpose of this study, an average lifetime in the order of a few days for all the investigated species is a reasonable statement. We will add a reference accordingly (Seinfeld and Pandis, 2016).

L137: please add: "… parts or all of the winter snow cover…"

Will be added.

L138: add per year after w.eq.

Will be added.

L139: Please weaken this statement to "… ice at bedrock can be more than…" and cite additionally: Hoffmann, H., Preunkert, S., Legrand, M., Leinfelder, D., Bohleber, P., Friedrich, R., & Wagenbach, D. (2018). A New Sample Preparation System for Micro-14C Dating of Glacier Ice with a First Application to a High Alpine Ice Core from Colle Gnifetti (Switzerland). Radiocarbon, 60(2), 517-533. doi:10.1017/RDC.2017.99

Statement will be weakened (*"and the observation that up to 15000 years old ice was found at the bedrock."*) and the suggested reference added.

L173: were the IC measurements discrete or quasi-continuous? What was the spatial resolution? Please specify.

*"on discrete samples"* will be added. The spatial resolution was different from core to core and is detailed in the different references given in table 1 and 2.

L174: Is there a reference for the IUP measurements?

The given reference (Preunkert et al., 2001a) does also include the IUP measurements.

Section 2.2.3: Are there no measurements available for the other cores? Please explain.

For this study we used records that were published until beginning of 2022. There are indeed no other published continuous measurements.

L197f: Is this sea salt and mineral dust correction valid for such rather continental sites?

Yes it is based on the evidence in the cited paper (Schwikowski et al., 1999b).

L201: It is unclear what happened after the separate calculation of summer and winter mean concentrations. Were they also evaluated separately? Please extend and clarify.

Based on the evidence given in the previous CDD papers, winter layers at the CDD site were found to generally thin with depth relative to summer layers due to changes in depositional processes upstream of the drilling site (Preunkert et al., 2000). Therefore annual values were defined as arithmetic mean of winter and summer concentrations. The separate evaluation of the summer and winter concentration records is not subject of the paper, since such a separation of summer/winter layers could not be done on other cores due to the low resolution in the pre-industrial parts. For the CDD it has been discussed in many previous studies (e.g. Preunkert et al., 2003).

L203: See general comments. How were the 10-year averages calculated? Just arithmetic mean? Weighted? Please extend and explain.

Arithmetic means were calculated, will be added.

L205: see general comments "should better reflect" is no sufficient justification for the 10-year averaging. Please assess better and explain reasoning.

See detailed answer to this general comment above.

L210: The reason for not including the GG records does not convince me. 50 years out of 250 is already 20%. If this is considered too short, the GG core should have been left out completely.

For calculating z-scores, averages and stdev of a common overlapping period have to be considered. We consider the 50 years period covered by the GG core too short (just five data points) to obtain representative values of these two measures. Following the referees comment for L282, we prefer to leave the GG records in Fig. 2 and related discussion in the manuscript to compare concentration data from this high-accumulation site (~2.7 m weq/yr) with the ones from the close-by low-accumulation CG site (~0.3 m weq/yr).

L211: Again, are there no BC /EC records available for the other cores? Clarify.

There are no other published continuous records.

L219 and Fig. 1: It is really hard to see significant differences between the FLEXPART results for the different locations. Maybe this figure could benefit from a larger zoom on the locations and / or a different colour scheme if the goal is to show the very small large-scale variations?

There are no substantial differences in the pollutant source areas for the different sites (line 246). Since the goal here is to show common general source areas in Western Europe, we prefer to leave Fig 1 as it is.

L221: "Most species" – which ones are not present as sub-micrometer aerosols? Specify.

We will change "*most species*" to "*all species are predominantly present…*"

L229: how exactly was this calculated? Please elaborate.

A detailed description of the calculation of the source-receptor relationship for deposited mass in FLEXPART is provided by Eckhardt et al. (2017). We will add a corresponding sentence at the end of the paragraph.

L234: spatial distribution referring to what region? Please specify

The MATCH model was applied to a European domain for both datasets (MATCH-BIODIV and MATCH-ECLAIRE), as indicated in the opening sentence of section 2.5. We used these simulations but focusing on the spatial variation in the Alps in this study (e.g. Fig 3). We will clarify that the modelled spatial distribution of $exSO_4^{2-}$ concentration in the Alps, used in the present study, was extracted from these simulations.

L238: which climate simulations? The whole paragraph 2.5 is rather vague. Please specify.

The climate simulations (and emission data) used for the MATCH-ECLAIRE simulations are described in the Engardt et al. (2017) paper, while the MATCH-BIODIV will be described in a forthcoming paper and in other present work. It is outside the scope of the present paper to explain all details of the atmospheric chemistry simulations here, but we will update paragraph 2.5 with more details.

L243: how do you define long-term-trend? What would happen for a smaller / larger averaging window?

We will change this term to "long-term variations" and specify this in the revised version. The justification for the choice of the 10-year averages is given above.

L254f: Please reformulate this sentence. It seems odd if you write "could be explained" when you dismiss this possibility later in the text. Also the west-east gradient seems confusing in this context, when in Fig. 1 you mention that for CDD the source region is extended to the west.

We will reformulate the sentence: "*Potential reasons for the lower CDD concentrations are…*"

L255: Do you have estimates for the ratio of the summer / winter snow accumulation? This would be an important information.

There are no reliable measurements of these ratios available for the high-altitude sites.

L259 and Fig. 3: Does that 50 km MATCH-ECLAIRE model resolution actually make sense if the cores are only separated by ~100 km? Maybe focus on the BIODIV version and don't show the low resolution runs? The gradient only seems to be visible in the low resolution plots, this is not sufficiently discussed.
L262: Again, what is the benefit of the MATCH-ECLAIRE, if it underestimates the altitudes in such way? Please consider focusing this paragraph only on the high-res version.

The MATCH-ECLAIRE set up is unique in the sense that it provides consistent pan-European trends of atmospheric concentrations and deposition of major atmospheric constituents over the period 1900-2020. Due to computational constraints and uncertainties in the input data during the early period this set-up was operated on relatively coarse resolution.

The MATCH-BIODIV simulations, on the other hand, use more recent input data, available on higher spatial resolution, which enable more detailed analysis of the spatial variation of sulphur- and nitrogen containing species across central Europe, but it is only available from 1987 onwards.

The two datasets when used together allows for a more comprehensive understanding of both spatial and temporal variations. We will improve the discussion on the modelling data sets to reflect these considerations.

The fact that the gradient is only visible in the low resolution model (MATCH-ECLAIRE), but not in the high-res MATCH-BIODIV model is already explained in lines 260-271.

L272: Please reformulate the first sentence. Is this a conclusion from the fact that the west-east gradient was ruled out?

Yes! We will reformulate the sentence accordingly.

L282: Given the close spatial proximity but very different accumulation regimes: how do the CG and GG records compare? This would be important to know in the course of the data evaluation.

As described in 3.2.-3.4. and fig. 2 concentrations agree for CG, GG, and FH. We will add a discussion accordingly.

L284ff: In my opinion this last statement of the paragraph questions the representativity of one single record (at least in total concentration apart from trend) for the region. I think this is contrasting one of the main messages of the paper. This needs further discussion and assessment.

In response to the previous referee comments we will clarify in chapter 3 that representativeness was investigated comparing a) absolute concentration levels and b) long-term concentration trends. Our conclusion is that every ice core from the different sites in the European Alps provides a representative signal of the trend of anthropogenic pollution in Western European countries. This does not hold for the magnitude in concentrations, for which CDD shows generally lower values,

mainly caused by a different seasonality in preserved precipitation. We will adjust this part of the conclusion, as suggested by the referee.

L297f: What about the CG15 peak at about 1920 in Fig. 2? This is also visible in NO3 and NH4, If it still shows up in the 10 year average, I would consider it significant enough to be discussed. Please comment.

The idea of the paper is not to discuss and compare single peaks but to focus on general features like concentration levels and trends. Therefore we do not discuss single 10-year maxima such as the CG15 1910-20 peak in $NH_4^+$, $NO_3^-$, $exSO_4^{2-}$, or the CG95 1880-1890 peak of Pb.

L326ff: Why should these uncertainties only be relevant in the late 20th century? There is almost a factor of 2 difference between observation and model in the modern section. This needs more critical discussion.

There are no differences between model and ice-core trends until end of 20th century. Potential factors for the deviation after this time are explained already (potential uncertainties in NOx emissions estimates or model misrepresentation of the $HNO_3/NO_3^-$ partitioning between gas and particle phase and of the atmospheric chemistry).

L339f: if analytical uncertainty is the reason, this is not sufficiently reflected in the uncertainty envelopes. Again, there are excursions for CG15 around 1910-1920 and FH ~1860. This needs a better assessment than the speculative comments. Please extend.

We agree that "analytical uncertainty" is not the correct term and will remove it from the discussion. What we referred to is a potential contamination from the laboratory air, which is not included in the uncertainty envelope. We will clarify that in the revised version.

L400: if the model emission reductions were too optimistic, could this be corrected and re-evaluated to find a better agreement? Please comments

The idea of the paper was to compare ice-core data with available model data. The re-evaluation of model data to fit the ice-core data is out of the scope of this paper and will be tackled in future studies.

L408: Typo: Methods of analysis

Will be changed

L408: you could add "to the CG cores" after "similar trend"

Will be added.

L408f: please clarify what you mean by "elevated Pb levels in the earlier…" does this refer to the relative change? the total concentration is lower anyway. The same applies for Cd.

Yes, it refers to the relative change, we will correct this for Pb and Cd.

L428f: If these are the locations of the smelters, why is there no larger impact in the CDD cores visible? Please comment

The distance of the French smelters to the CDD site is comparable to the distance of the Swiss smelters to the other three sites.

L442f: There is no sufficient answer given to the question of representativity throughout the manuscript. This needs more specific discussion.
L449f: This representativity is in the current stage of the study only true for the long-term trends, but not for the absolute concentrations. This needs clarification.
L455ff: Does this mean that CDD is less representative for the region than the other cores? This also needs more critical assessment and discussion.

See answers to general comments and to line284.

L458-462: This paragraph also needs some evaluation .It only lists the different TOE. What are the consequences for the present day situation?

The present day situation is discussed in the following paragraph (L464-471)

L476-L479: Please reformulate this final statement. It is common knowledge that beyond the instrumental era ice cores are the main tool to reconstruct aerosol deposition. What especially is the knowledge gain emerging from this study? Is it representativity of the single records? Or rather pinning down the onsets of pollution? This needs clarification (see also general comment).

In our study we show for the first time, how consistent the Western European air pollution trends are recorded in single Alpine ice cores. We were able to refine the history for different compounds by pinning down their onset of pollution, maxima, and recent changes. The fact that only four of seven pollutant records do agree with available model estimates illustrates the necessity to include such ice core-based reconstructions for model constrain. We will adapt the final statement accordingly.

Table 1: It would help to add a column with the core length and / or estimated maximum age at the bottom.

Since we use only a part of the respective ice cores covering the period 1750-2015, we think the total core length as well as the maximum age at the bottom is not relevant. This information is available in the given references.

**References:**

Bohleber, P., Wagenbach, D., Schöner, W., & Böhm, R. (2013). To what extent do water isotope records from low accumulation Alpine ice cores reproduce instrumental temperature series? Tellus B: Chemical and Physical Meteorology, 65.

Eckhardt, S., Cassiani, M., Evangeliou, N., Sollum, E., Pisso, I., and Stohl, A.: Source-receptor matrix calculation for deposited mass with the Lagrangian particle dispersion model FLEXPART v10.2 in backward mode, Geosci. Model Dev., 10, 4605-4618, 10.5194/gmd-10-4605-2017, 2017.

Engardt, M., Simpson, D., Schwikowski, M., and Granat, L.: Deposition of sulphur and nitrogen in Europe 1900-2050. Model calculations and comparison to historical observations, Tellus B, 69, 10.1080/16000889.2017.1328945, 2017.

Preunkert, S., Wagenbach, D., and Legrand, M.: A seasonally resolved alpine ice core record of nitrate: Comparison with anthropogenic inventories and estimation of preindustrial emissions of NO in Europe, J. Geophys. Res.-Atmos., 108, 4681, 2003.

Schwikowski, M., Döscher, A., Gäggeler, H. W., and Schotterer, U.: Anthropogenic versus natural sources of atmospheric sulphate from an Alpine ice core, Tellus B, 51, 938-951, 10.3402/tellusb.v51i5.16506, 1999b.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric Chemistry and Physics: From Air Pollution to Climate Change, John Wiley & Sons, Hoboken 2016.

Sierra Hernandez, M.R.,  P. Gabrielli, E. Beaudon, A. Wegner, L. G. Thompson, Atmospheric depositions of natural and anthropogenic trace elements on the Guliya ice cap (northwestern Tibetan Plateau) during the last 340 years. Atmos. Environ. 176, 91–102 (2018).