

Reviewer 1

This is a clearly written paper with a controversial conclusion. The fact that temporal trends in the two polar regions are significantly different challenge our understanding on either atmospheric process or the presented firn modeling efforts. This is adequately discussed in the manuscript. Hopefully, the paper will stimulate further research to solve this unexpected result. Presentation quality, language and level of scientific argumentation are excellent. The paper should be published in ACP

We thank the reviewer for his supportive comments.

Reviewer 2

Patterson and co-authors present new firn air-based atmospheric reconstructions of atmospheric H₂ over the last century, combining data from several sites and both hemispheres. The difficulty in reconstructing atmospheric H₂ from firn and ice core records is the molecular size-dependent bubble close-off fractionation. This fractionation results in H₂ depleting in mature ice, and a corresponding enrichment in the firn air. The authors present several ways of correcting for this firn artifact.

I find their reconstructions convincing and worthy of publication in *Climate of the Past*. The work seems technically sound, and has been described clearly. I have a few suggestions I would like the authors to consider, in particular regarding the structure of the manuscript.

We thank the reviewer for his supportive comments, thorough review, and detailed suggestions. Specific comments are addressed individually below.

(1), I wonder why the authors chose to present the “regular” UCI_2 inversions as their main result (Figs 1-4), and then the various scenarios with altered pore close-off schemes as some sort of alternate or special case (Figs. 5-9). In my mind, one cannot reconstruct atmospheric H₂ meaningfully without getting the close-off size fractionation correctly, and therefore I think the “altered” scenarios are far more convincing. Case in point, these “altered” scenarios provide a good fit to the atmospheric flask observations that are the gold standard in atmospheric trace gas reconstruction.

If this were my paper, I would not bother with the regular UCI_2 model, as it is clear it cannot fit the Neon data nor the direct atmospheric flask data. It is customary to calibrate (or tune) the diffusivity profile to trace gas records; in this situation I would consider fitting the Neon data as part of the model calibration (tuning). I would have only presented the “altered” scenarios. This would also result in a much shorter and more focused manuscript. I suspect the authors trust the altered scenarios better themselves, as this is what they plot in their last summary figure (Fig. 9). Perhaps this review can be a justification for the authors to make this simplification of the manuscript. On the other hand, I realize this is a big revision I am suggesting, and I would still be supportive of publishing this paper if the authors decide to keep the current structure.

At the very least the authors should communicate more clearly which inversion they believe to be most realistic, so that users of these reconstructions know which one to plot. I would advocate strongly that the most realistic ones are those that can fit the dNe/N₂.

We agree with the reviewer that the “altered” scenarios are more realistic and that this should be clearly communicated in the paper. We prefer to include the UCI_2 model results in the manuscript as a “baseline,” in order to contextualize need for altered scenarios, particularly since the altered scenarios are essentially *ad hoc*. Including the UCI_2 model results also demonstrates the dramatic differences between the Greenland sites and the Antarctic sites.

Revisions related to this point:

L751-753: Absent additional measurements, we believe that the results presented in Figure 9 are the best estimate for the evolution (and uncertainty) of atmospheric H₂ during the 20th century over Antarctica and Greenland.

L139-154: The UCI_2 firn air model is a 1-dimensional finite-difference advective-diffusive model that is used to simulate the evolution of trace gas levels in firn air. The UCI_2 model has been used to successfully analyse H₂ levels in firn air at two Antarctic sites, including the effects of pore close-off fractionation (see below; Patterson et al., 2020; 2021). As discussed in Section 7, the default model parameterizations do not adequately capture the effects of pore close-off fractionation at two of the Greenland sites; nonetheless, the model provides a useful baseline for analysing atmospheric H₂ levels in firn air.

L746-748: Additionally, the alternative parameterizations of pore close-off fractionation yield better model-measurement agreement for $\delta^{22}\text{Ne}/\text{N}_2$. The improved agreement indicates that the reconstructions generated by the alternative parameterizations are more realistic than the reconstruction generated by the UCI_2 model.

(2) The authors apply a gravitational fractionation correction to the data via the d¹⁵N data. It would be trivial to similarly apply a close-off fractionation to the data via the dNe/N₂ data (assuming that the dH₂/N₂ fractionation is the same as the dNe/N₂, which the authors assume already). In this way the bubble close-off would be included empirically, and would not have to be modeled. Can you add an inversion using such an empirical data correction? My prediction would be that it matches the “altered” solutions.

The inversion described above is not identical to the alternative solutions. This is due to the small difference in the age distribution of H₂ caused by mixing between the closed bubbles and open pores. As a result, explicit modeling of pore close-off fractionation is necessary. We have added a Figure to the supplement demonstrating this (Figure S7), and made the following addition:

L495-497: Because mixing between the closed bubbles and open pores affects the modeled H₂ age distributions, measurements cannot be empirically corrected for pore close-off fractionation as they are for gravitational fractionation (equation 2; Figure S7). Instead, pore close-off fractionation must be explicitly modeled.

(3) The discussion around the close off fractionation (section 7) is not as clear as it could have been. A few suggestions:

(3a) Can you add a panel to Fig. 7 showing the bubble pressure in the model? That is ultimately what drives the fractionation, so it is critical to have this information.

We like the reviewer' suggestion to show the physical factors responsible for fractionation on Fig. 7. At equilibrium, modelled pore close-off fractionation is controlled primarily by the ratio of the volume-weighted average pressure in the open and closed pores to the ambient pressure, adjusted for a small amount of mixing. We decided to plot this ratio (R) rather than bubble pressure alone, making it more obvious for readers why the three alternative parameterizations yield the same neon enrichment. The text was also revised as follows:

L468-481: At equilibrium, modelled pore close-off fractionation is controlled primarily by the ratio of the volume-weighted average pressure in the open and closed pores to the ambient pressure, adjusted for mixing. We introduce a new parameter, R :

$$R = (P_{bubble}S_c/S_{total} + P_{ambient}S_o/S_{total})/P_{ambient} \quad (12)$$

In the UCI_2 model, R begins to increase too deep in the firn to capture the shallower $\delta^{22}\text{Ne}/\text{N}_2$ measurements (Figure 5; Figure 7). As a first attempt to improve agreement between measured and modeled Ne enrichment, we re-tuned the closed porosity profiles for NEEM and Summit to optimize model-measurement agreement for Ne enrichment using previously published parameterizations, including Schwander et al. (1989), Goujon et al. (2003), and Mitchell et al. (2015). However, in all cases, the optimization yielded closed porosity profiles that are qualitatively different from previously published estimates and probably unrealistic. When the optimizations were constrained to generate realistic closed porosity profiles, in all cases, we found that R increases too rapidly with depth capture the $\delta^{22}\text{Ne}/\text{N}_2$ at both the top and the bottom of the lock-in zone. Therefore, in order to generate the necessary R profile while maintaining a realistic closed porosity profile, some other physical process must be modified in addition to the closed porosity profile.

(3b) I don't understand why the "reduced compression" scenario would result in more dNe/N_2 fractionation. That makes no sense physically, as halving the pressure should also half the dNe/N_2 anomaly. From Fig. 7a I suspect that the authors also altered the closed porosity parameterization, and that this is what drives the enhanced dNe/N_2 . Please check/confirm.

The reviewer is correct. The closed porosity profile was changed in addition to the compression. We have tried to clarify that in all three scenarios, the closed porosity parameterization is changed in addition to some other physical process (see response to 3a). In addition, L571-572 now states: "Again, the original closed porosity profile does not yield agreement with the measured Ne enrichment, so a new closed porosity profile is prescribed using equation 12. The three tuneable parameters from equations 12 and 13 (k_{comp} , m , and ρ_0) are simultaneously optimized to give the best fit between modeled and measured Ne enrichment (Table 3; Figures 5 and 7)."

(3c) I am not surprised that the regular NEEM closed porosity parameterization gives a poor result for dNe/N_2 , given that it closes off much too deep (resulting in less pressurization). The Goujon/Martinerie close-off density at that site was artificially increased to match the field observation of the deepest pumping depth. It also ignores layering; including layering automatically results in some shallower trapping that will increase the dNe/N_2 anomaly (as shown in Fig. 7a). Have you tried using the porosity parameterization from Mitchell et al. (2015), with the suggested close-off density from that paper? That

may solve some of these problems. That parameterization does not produce an abrupt full bubble closure as may be required in some model architectures – this can be added manually perhaps.

We did try the Mitchell parameterization when trying to tune the model to capture the neon enrichment, but the results were not optimal. This is now mentioned in the text (see response to 3a).

(3d) The solutions from all three “altered” scenarios are virtually identical. Could this be because they start trapping bubbles at the same depth (Fig. 7)?

Essentially, yes. More specifically, the three alternative scenarios yield the same neon enrichment because R is essentially identical in all 3 cases (see revised Fig. 7).

(3e) Could you extend the plot in Fig. 7 further down, to for example 90 m? Currently we cannot evaluate how sudden or deep the full bubble close-off occurs (the point where the closed porosity starts to decrease). Sometimes I find it more useful to plot the closed pore fraction, rather than the close porosity itself.

Our model domain does not extend deeper than the depth where closed porosity fraction=1 (see revised Fig. 7). This means that for the UCI_2 model, we have plotted to the point where closed porosity begins to decrease. Because the alternative scenarios are ad hoc, we do not think that extending the domain adds scientific value to the paper. New high resolution measurements of firm density, porosity, and bubble total air content are needed to investigate this question.

(4) Can you elaborate on using a Green’s function approach in the presence of bubble close off fractionation? Green’s functions assume a linear system response (the sum of two solutions is also a solution to the diffusion equation). Is this true in the presence of close-off fractionation? I suspect it is, but I am not entirely sure. Is the area of under the Green’s function greater than 1 in this case?

The reviewer is correct on all accounts. We have tested this by comparing forward model runs to equation 7. We have made the following addition to the manuscript:

L241-243: For most gases, the Green’s functions sum to 1 at every depth. In the case of H₂ and Ne, the sum of the Green’s functions in the lock-in zone is >1 due to pore close-off fractionation.

Other comments:

L33: “second-most”

Done

L41: consider replacing “modern” with “present-day”

Done

L45: Do you have a reference for the OH radical?

The references at the end of the sentence include estimates of the lifetime of H₂ with respect to oxidation by OH. Those papers include information about the OH fields used in their estimates.

L52: Consider also adding a reference to Solomon et al. 2010, who first clearly describe greenhouse forcing from stratospheric H₂O

Done

L62: These trends are not very robust, and rely on single year anomalies. Also, how well are Khalil and Rasmussen calibrated with NOAA/GML?

Khalil & Rasmussen data show an increasing trend from 1986-1989 and the NOAA data show a decrease from 1989-1993. Note, there is a more detailed discussion of the flask data in L387-394. We have made the following revisions:

L56-57: Integration of records produced by the different groups has been complicated by calibration issues, which are discussed in greater detail in Section 5.

L61-63: Broadly, the instrumental record shows northern hemispheric H₂ levels rising during the late 1980's to a maximum in 1990 and decreasing until 1993. There is no discernible trend in Northern Hemisphere H₂ levels from 1993-2010 (Figure 3).

L79-84: What about Antarctica? Those SH reconstructions are treated as somewhat of an afterthought in this paper, despite the topic being bipolar H₂.

This is true... the Antarctic reconstructions are not discussed in as much detail because the differences from our previous work (Patterson et al., 2020; 2021) are small. We have made the following addition to the manuscript:

L81: Additionally, we reanalyse the Antarctic firn air data using a different inversion technique, described in Section 3.2.

Section 2.2: Normally the diffusivity profile is somewhat model-dependent. Can you simply apply the profile from a different model?

We validated the diffusivity profiles by simulating CO₂, CH₄ (Figure S8). We also ran SF₆ and CH₃CCl₃. See L193-197.

L134: same depth "were" averaged

Done

L136: intense seasonality: is it much deeper for H₂ than for other gases like CO₂?

Yes, for two reasons:

1) The seasonality of H₂ in the NH is ~10%. For comparison, the seasonality of CO₂ is ~4% and the seasonality of CH₄ is ~2%

2) The free air diffusivity of H₂ is much higher than that of other gases, so seasonal levels penetrate deeper into the firn

L144: different sampling dates?

Text revised to clarify that the differences in profiles reflect both differences in site physical characteristics and changes in atmospheric H₂ over the time period during which the various firn air studies were done (17 years):

L140: The qualitative differences in the depth profiles reflect changes in surface air H₂ levels over the period of the firn air studies (1996-2013) and the different physical characteristics of the sites.

L154: This upper 5m is often called the convective zone

We prefer to leave it as two zones because the mathematical/coding treatment of the convective zone and diffusive zone is identical.

L154: this is often called wind pumping or just ventilation. Convection happens in winter when the surface is cold.

We have removed the references to convection.

L158: that is a very small time step! I typically run my firn air model with a timestep of one week or so. Is this needed to keep the forward Euler scheme stable?

Yes, the small timestep is necessary because of the forward Euler integration scheme and the high diffusivity of H₂. The model can be run at a higher timestep for other gases.

L170: Which parameterization? Goujon? Schwander? Mitchell?

We used Schwander for all three sites. We have added a footnote to Table 1 (L105).

L175: The δ¹⁵N data also have thermal fractionation in them. How do you deal with this?

Thermal fractionation is not important below the upper part of the firn, and we exclude those measurements from the reconstruction due to seasonality. We have made the following revision:

L1835-187: This correction neglects the thermal fractionation of δ¹⁵N. Thermal fractionation is only important in the upper ~20 m of the firn, and these shallow measurements are excluded from the reconstructions due to seasonality (Severinghaus et al., 2001; Section 3.2).

L189: Confirm that the model is coded in volumetric concentrations, rather than in ppm. Most models work in ppm, I believe.

The model is coded in moles/m³ as noted in the text.

L197: gas phase diffusivity: do you mean free air diffusivity?

Yes, fixed.

L202: This is not really eddy mixing of course, though mathematically it is similar. This is more correctly described as dispersive mixing (Buizert and Severinghaus, 2016)

We have revised the text accordingly:

L205-207: At the top of lock-in, small, non-fractionating values of “eddy diffusivity” are prescribed to account for dispersive mixing caused by barometric pressure fluctuations (Buizert & Severinghaus, 2016).

L201-206: so the advection is coded differently in the diffusive and lock-in zones? Is it a velocity term in the former, and a box-shuffling scheme in the latter? Does this conserve mass at the boundary? Do you account for the fact that there is backflow in the lock-in zone due to compaction?

The reviewer is correct about velocity in the diffusive zone and box-shuffling in lock-in. The box shuffling prevents numerical diffusion in lock-in. To couple the two schemes, we use a “buffer” as described in the supplement of Severinghaus et al., 2010. We are impressed that the reviewer thought to ask about backflow. We neglect the backflow term simply because turning it off yields a better empirical fit to tracers with well-constrained atmospheric histories such as CO₂ and CH₄. We have made the following revisions to the text:

L154-164: The UCI_2 model is largely based on Severinghaus et al. (2010). The model domain is divided into an upper “diffusive zone” and lower “lock-in zone.” In the diffusive zone, vertical gas transport occurs via wind-driven mixing in the shallowest ~5 m and via molecular diffusion throughout. Diffusive mixing decreases with depth due to the increasing tortuosity of the firn. In the lock-in zone, vertical molecular diffusion ceases due to the presence of impermeable ice layers. Gas transport in the lock-in zone occurs primarily due to advection with a small non-fractionating mixing term. The model uses a forward Euler integration scheme and a time step of 324 s., There are three important differences between the Severinghaus et al. (2010) model and the UCI_2 model: 1) thermal diffusion is neglected, as it is unimportant for H₂, 2) backflow due to densification in the lock-in zone is neglected in order to yield a better empirical fit to the tuning gases, and 2) our model parameterizes pore close-off fractionation differently than the Severinghaus model (see below). The model tracks the air content and composition in both open pores and closed bubbles as a function of time and depth. The model code is written and executed in MATLAB R2022a (Mathworks Inc.).

Equations 4-6: How is this implemented? The x_n and P_{bubble} terms occur in all three equations, so you cannot simply solve them. Is this done iteratively? Or is there a typo in the equations?

P_{bubble} is a function of depth, but invariant in time and P_{ambient} is constant. At the end of each time step, equations 4-6 are solved sequentially. No iteration is required.

L235: parenthesis) missing after Rommelaere citation

Fixed

Equation 7: Can you add the arguments to the variables here to clarify? For example, $G(z,t)$ etc

Yes, done

Eq 8: What is N ? Normal distribution? What is the time step i ? 1 year?

The reviewer is correct: We have made the following revisions

L259-260: Where N is the normal distribution, and m_{atm} is a vector of length i that contains the discrete atmospheric H₂ dry air mole fraction history with a timestep of 1 year

L267: There is no artificial smoothing, but instead the parameter beta in the autoregression. Isn't that just the same with a different name?

There difference is that we don't arbitrarily prescribe the amount of smoothing, as done in the Rommelaere method. We have made the following revision:

L272-273: 1) No arbitrarily specified smoothing criteria is imposed (instead the atmospheric history is assumed to be autocorrelated)

L275: With such cut-off depths, you have only 1 to 8 m in the diffusive zones. Can you confirm?

Confirmed

L 275: Can one instead add a seasonal cycle to the atmospheric history/ inversion?

The answer to this question is a little complicated. Because of the necessity of auto-correlation (or smoothing), we would need to superimpose the seasonal cycle on the long-term trend. This would only provide additional information about the year before the firn air would sample. Those times are better constrained by atmospheric observations, so this effort would provide little scientific value.

Figures 1 and 2: is it possible to plot the firn data on the plot against their mean age?

Yes, we have added it.

L354: Can you add the Alert data to Fig. 3?

The Alert data is in Fig 3 (yellow), but we did catch a typo in the legend.

L357: Do you have a reference there for the ENSO connection? What about 1989?

We have added references for the ENSO connection and made the following addition to the text:

L392-395: The reason for the rapid decrease in observed H₂ levels in the NOAA/GML data is not clear. At that time, anthropogenic emissions of H₂ were likely decreasing but not rapidly enough to account for the observed decrease (Hoesly et al., 2018; Paulot et al., 2021). It is possible that the observed decrease is linked to NOAA's drifting calibration scale as discussed in Section 2.2.

L397: suggested that "the" maximum...

Fixed

L409-410: But the maximum mostly disappears when you account for the close-off fractionation

Yes, we have made the following revision:

L433-438: The late 20th century maximum in atmospheric H₂ is a robust feature of these Greenland firn air reconstructions. According to the UCI_2 model, pore close-off induced enrichment is <1% at the depth of the observed H₂ maximum in the lock-in zone. That is, the observed lock-in maximum is not caused by enrichment, but by a historical atmospheric maximum. However, there is evidence that the UCI_2 model underestimates the impacts of pore close-off fractionation at the Greenland sites. The implications of this underestimate for the atmospheric maximum are investigated and discussed in Section 7.

Fig. 4: is it possible to plot the firn air data against their mean (or even effective) ages on the figure? Possibly empirically corrected for close-off fractionation? I always find this extremely helpful, as it allows

the reader to visualize the data density and the degree of smoothing. As an example of plotting in this style, look at the recent Ghosh et al. (2023, <https://doi.org/10.1029/2022JD038281>)

Yes, done

L468: Have you also compared to Mitchell et al. 2015? The Goujon and Schwander parameterizations do not account for density layering, and are commonly applied incorrectly (i.e., they were derived on cm-scale hand-samples and are applied to m-scale bulk density; because of density layering this is technically incorrect. Layering broadens the depth of bubble trapping).

See response to 3 above.

Fig. 5: It seems clear that the UCI_2 model does not build up pressure fast enough in the pores to expel fugitive gases. Shallower trapping seems like the obvious explanation to me, and tuning the closed-porosity parameterization makes sense to me as the first strategy. The Mitchell et al. 2015 parameterizations has two parameters that can be tuned.

See response to 3 above.

Fig. 5: can you also show the fit to the Antarctic dNe/N₂ data somewhere? These sites are used, and so the reader will wonder whether the modeling can fit those data.

Yes, we have added it to the supplement (Figure S8) and referenced it in the text at L484.

L520+: can you give a plot of the s_c parameter? This is hard to visualize.

We have plotted s_c/s_{tot} in Figure 7. We prefer not to plot both because of the similarity.

Eq 13: you have different numerical advection schemes in the diffusion and lock-in zones, right? How does this impact the implementation of the bubble pressurization?

It does not. Bubble pressure is constant in each layer in time. It is calculated using quantities (new bubble volume and closed porosity) that are normalized against the volume of each layer, so the different layer thicknesses do not matter.

L547: You seek to explain the effect via the bubble compression rate only, but then alter the porosity profile after all to get something similar to the first scenario. It is the porosity tuning that makes you fit the data, not the reduced bubble compression rate – if anything the latter should make it harder to fit the data. I think this is confusing to the reader. Why not instead conclude that solely altering the pressurization rate (which this scenario ostensibly represents) does not improve the fit?

We have tried to clarify our treatment here. See response to 3 above.

Fig. 7: can you add plots of the NEEM and Summit closed pore pressures? That seems needed to evaluate the fugitive gas enrichment.

See response to 3a

L678: what about H₂ artifacts during flask storage?

We can only reasonably constrain these at NEEM due to a lack of data from other sites. For NEEM, the effects were small (~1%). In lieu of trying to constrain the blank, we increased the uncertainty on the firn air measurements by a factor of 10 (see discussion in L126-129)