Reply to Referee 2

Thank you for your constructive and detailed review of the manuscript. Our response to your comments and the changes we plan to make to the manuscript are annotated below. The text from the review comments is in black italic text, our responses are in blue, and the changes we intend to make to the manuscript are in red.

Anonymous Referee #2
Referee comment on "Gas isotope thermometry in the South Pole and Dome Fuji Ice Cores provides evidence for seasonal rectification of ice core gas records" by Jacob Davies Morgan et al., The Cryosphere Discuss., https://doi.org/10.5194/tc-2022-49-RC2, 2022

This manuscript deals with paleothermometry based on new measurements of gas isotopes in the South Pole ice core. Some measurements from the Dome F ice core are also presented. The authors are presenting a very thorough description of the methods and present improvements in the precision of d15N and d40Ar measurements which is impressive. Using firn densification modeling combined with the series of measurements of d15N and d40Ar over the South Pole ice core, they propose a reconstruction of the firn thickness and temperature gradient between the top and the bottom of the firn over the period covering 30 to 5 ka. The interpretation of the variation of the temperature gradient in the firn is not easy. Several scenarios are proposed and the authors conclude with the existence of a seasonal bias affecting the gas isotopes record.
This manuscript is well written and details the different steps of the methods and of the reasoning. It should be published within TC. I still have several comments which I think should be addressed before publications.

General comments:
I suggest to remove the whole section focused on Dome F. It is a bit disconnected from the study of the SP DTz and DCH. This section is also difficult to follow since it is not enough documented (the d15N and d40Ar data are not shown nor the origin and associated uncertainty for the DTz modelled curve). Moreover, if the Dome F data are shown, we may also wonder why we can not have the same for other sites ? showing or not a seasonal rectifier effect.

Because this comment is closely related to your final comment, we address the two together at the end of this document. We made significant changes to Section 5.2.4, including expanding our explanation of the Dome F data so that their relevance to the study is clearer to the reader.

The results displayed here raises doubts on the classical interpretation of d15N-excess (Kobashi et al., 2007; Kobashi et al., 2011) in term of surface temperature variations. A discussion revisiting these previous studies should be included here as well as clear recommendations on how to use or not the d15Nexcess for reconstructing past surface temperature variations.

We will add a paragraph to the "Broader Implications" section that discusses the relevance of our work to previous $\delta^{15}N_{excess}$ studies.

Rectification of ice core gas records has received limited attention in the literature so far, but our work argues that more careful consideration is necessary. Failure to recognise and account for rectifier effects where they do exist could potentially lead to incorrect temperature estimates. Fortunately, it is unlikely that rectifier effects would have been significant for previous gas isotope thermometry studies in Greenland (e.g., Kobashi et al., 2007, 2011; Orsi et al., 2014; Landais et al., 2004, 2006; Huber et al., 2006). The presence of rectification via the mechanism we describe likely requires specific surface conditions such as stagnant air and a strong atmospheric temperature inversion. These conditions probably occur rarely on the Antarctic plateau and are even less common in Greenland. To have any effect on the composition of air in the deep firn and closed-off ice they must reoccur every year for many decades. Furthermore, in the case of Kobashi et al. (2011), agreement between their temperature reconstruction, regional climate model outputs and modern instrumental records also supports their analysis and interpretation. However, it might be necessary to account for rectifier effects in future gas isotope thermometry studies in Antarctica.

Comments along the manuscript:
l-25 : Precise which « temperature difference » you are speaking about.

We state on line 20 that the temperature difference is between the top and bottom of the firn column. This seems like enough detail for the abstract. The details are explained in the body of the text (section 2).

l-41: The authors were aware that the spatial temperature – isotope relationship was a surrogate for the temporal relationship and always tried to check if this was true. So I suggest to replace "thought" by "assumed"

Good point. We will rewrite the sentence as below.

Early studies calibrated $\delta^{18}O_{ice}$ using its modern-day spatial relationship with mean annual temperature near the ice core site, which was hypothesized to be identical to the relationship with temporal variations in site temperature

l-136: I do not understand why the 30 minutes delay is important for the reference gas only ? Should it not be also the case for the sample gas ?

The cooling of the bellows only affects the reference gas because it is at a higher initial pressure than the sample gas, so there is a greater decrease in pressure and therefore a greater amount of adiabatic cooling (see Section S1). However, this sentence was worded in a confusing way because the delay happens after both sample and reference gases have been admitted into the mass spectrometer. We do it this way so that both experience the delay equally. This reduces the chance that we inadvertently introduce an additional bias by treating the sample and reference gases differently. We will rewrite to include reference to the pressure difference between the sample and reference gases and to make it clear that the delay happens after both the sample and reference gases have been admitted.

The second is the inclusion of a 30-minute delay between admission of the sample and reference gas into the bellows and the beginning of the measurement sequence. This is necessary due to an initial measurement bias caused by cooling of the bellows during expansion of the reference gas, which is at a higher pressure than the sample gas prior to expansion. Both modifications are discussed in more detail in Sect. S1.

Table 1 and associated text: I am not sure how relevant it is to compare the d15N results between "Orsi Ice" and "This study SPC". Indeed, "Orsi" and "This study" obtain the same results on air measurements and the improvements mentioned in the methods section apply on both air and ice. What could explain a better precision only for ice then ?
We can thus wonder if the difference is not simply due to a poorest ice quality in "Orsi" ?
Also, it should be noted that the replicates number by "Orsi" is much larger (169) than for this study (14) which makes the comparison questionable. Can you also provide the numbers of replicate fo the LJA analyses by Orsi ? By the way, given the length of the record presented in Figure 1, I am surprised that Table 1 presents only 14 replicates. This

should be better explained. It would also be useful to give the number of replicates for Kobashi's data.

We added the number of LJA replicates from Orsi (2013) and the number or LJA and Ice replicates from Kobashi et al. (2008). Thank you for pointing out that information was missing. Orsi (2013) do not report a pooled standard deviation for LJA $\delta^{15}N_{excess}$ and Kobashi et al. (2008) does not report a pooled standard deviation for LJA or Ice $\delta^{15}N_{excess}$. We also made several other changes to the table based on comments from another referee (see below).

For this study, we were more severely limited by sample size due to the smaller diameter of SPICEcore relative to previous cores. This made it very difficult to make as many duplicate measurements as we would have liked. However, we worked hard to make sure we were able to analyze 14 duplicate samples to give us some estimate of the reproducibility.

You are correct that it is possible the improved precision is due to better ice quality for SPICEcore compared to WDC. We will add this comment to Section 4.1.

We also note smaller improvements in the reproducibility of the other measurements and that some of the improvement may be due to superior ice quality for SPICEcore.

Table 1. Mass normalised pooled standard deviation of replicate measurements of $\delta^{15}N$, $\delta^{40}Ar$, $\delta Ar/N_{2\ grav}$, and $\delta^{15}N_{excess}$ from either reference gas runs (REF), La Jolla air flasks (LJA), South Pole ice core samples (SPC) or other ice core samples. Units for all four isotope ratios are ‰ amu^{-1} and the mass differences are 1, 4, 12, and 1 amu respectively. The final column indicates *n*, the number of samples used in the calculation.

	$\delta^{15}N$	$\delta^{40}Ar$	$\delta Ar/N_{2\ grav}$	$\delta^{15}N_{excess}$	Num. Replicates
This Study Ref	0.0020	0.0023	0.0080	0.0023	58
This Study LJA	0.0027	0.0024	0.0042	0.0019	40
This Study SPC	0.0022	0.0030	0.0432	0.0013	14
Orsi LJA	0.003	0.0025	0.0073		10
Orsi Ice	0.005	0.0036	0.0331	0.0042	169
Kobashi LJA	0.004	0.0035	0.0114		
Kobashi Ice	0.004	0.0040	0.0442		

Section 5.2 (first paragraph): The arguments developed here are a bit complicated to follow after the previous section where you explained that DCH is controlled by accumulation rate itself influenced by topography. And here, you say that we expect a link between accumulation rate and temperature. I thus suggest to rewrite this paragraph so that it is coherent with the findings of the previous section.

We rewrote this section to explain our logic more clearly and to link better with the previous section.

The variability in our record of ΔT_z is initially challenging to explain. We would have anticipated a positive correlation between DCH and ΔT_z since an increase in the accumulation rate ought to result in a thicker firn column and a weaker influence of geothermal heat on the temperature at the lock-in depth. However, although DCH and ΔT_z both increase over the course of the deglaciation, we instead observe a negative correlation between DCH and ΔT_z throughout most of the record (Figure 2). There must be some other mechanism that links variability in ΔT_z to either changes in accumulation or the local topography.

Section 5.2.1 – you mention that you are using the Dage to make the reconstruction of temperature and accumulation rate but the Dage model – data fit is not shown (nor any Dage data) and it is thus difficult to follow this discussion. Moreover, when looking at the

DTz (REF), it seems that the shape of the record does mainly depend on the d18Oice – can you explain better this reconstruction of temperature ? It is important to show which data are used to constrain the shape of the temperature evolution when DECOUPLE and REF disagree. I also expect that the shape of the Kahle reconstruction is mainly imposed by the water isotopes so actually it is expected that both Kahle and REF scenarios have the same shape. This resemblance should not be taken as a strong argument to discard the surface temperature influence on the DTz scenario.

You are right to point out that the Kahle and REF scenarios are not completely independent. Both use the empirical Δage data as a constraint. The Kahle reconstruction combines this with the diffusion length proxy to calibrate the water isotope record, so its shape is very much set by the shape of the water isotopes.
The REF and DECOUPLE scenarios both use the water isotope record as an initial estimate of the temperature. However, the final shape of the curve is not dictated by the initial estimate, but by the observational constraints applied during the optimization. This is clear from comparing the REF and DECOUPLE runs, which both use the same d18O data as the initial temperature template yet look completely different after optimization due to the different constraints. For the REF fun, the constraints are DCH and Δage. For the DECOUPLE run, the constraint is our ΔT_z reconstruction. Thus, even though both experiments start with the same initial estimate, the model can produce very different final temperature histories. Therefore, the resemblance between the Kahle reconstruction and the REF scenario comes partially from the fact that they both use empirical Δage as a constraint, and partially from the fact that the other constraints (DCH and diffusion length) are in good agreement with one another. We will rewrite the paragraph to explain this caveat more clearly.

Also shown for comparison is the optimal temperature history from the REF run (Buizert et al., 2021) and a temperature history from Kahle et al. (2021) based on a calibration of δ18Oice using the SPICEcore Δage data and the diffusion length of water isotopes in the firn. Note that both temperature histories are partially constrained by the Δage data, so they are not wholly independent.

Figure 6 – is there a way to add the DTz data so that the reader sees immediately that there is a mismatch

We tried a few different ways of overlaying the data or comparing the rate of change between the data and the modelled response to a change in GHF. However, we feel that adding extra lines or shading would make this plot more difficult to interpret. It is already a busy figure, so we think it is best to not add anything else.

l-424: the mechanism is not clearly explained – this part should be rewritten.

We expanded the explanation of the convection parameterization based on your comment and comments from the other referee. The re-written paragraph is included here and also in the response below this one.

In the model run without convection, the gases diffuse towards gravitational and thermal equilibrium as they are slowly advected downwards with the densifying firn and occluded in bubbles in the lock-in zone. Because the model is one-dimensional, it is not possible to explicitly simulate a three-dimensional Rayleigh-Bénard convection cell. Instead, we model just the sinking core of a convection cell, which we parameterise as an 8 cm d^{-1} downward transport of gas between 0 and 20 m. Between 20 and 25 m, the downward transport decays to zero, resulting in mass convergence that would be balanced in the real world by horizontal transport and a return flux of gas to the surface. This approach allows us to approximate how the gas isotopes respond to convection using a one-dimensional model.

- Figure 7: I am confused since the different data seem not 100% coherent with the provided explanations so probably more explanations are needed. If there is a temperature rectifier effect as suggested by the mismatch between model and data on the top 16 m, we expect a difference between d15N and d40Ar at the bottom of the firn which would then lead to a d15Nexcess signal due to seasonal rectification. Here, we see a difference but at 16m depth. Moreover, the d15Nexcess profile shown on the figures 1 and 3 does not show any 15Nexcess signal in the bubbles for the recent period, suggesting no difference between d15N and d40Ar at he bottom of present-day firn. Could the authors then better explain how they link their observation on the firn and the obesrvations in the air bubbles.

Thank you for pointing out some of the inconsistencies in the explanation of this figure. We have restructured the argument, expanded the explanation of the relevance of this figure to our argument, and added an additional figure that shows the rectifier in the deep firn. Hopefully it is clearer now, although we have not included the full re-written section here in order to keep this response somewhat concise. Briefly, we intend for Figure 7 to provide evidence that air convection in firn can advect thermal isotope signals deeper into the firn than diffusion alone can. Because the wintertime observations of δ^{15}N and δ^{40}Ar were made only between 0 and 16 m, it is possible they are the result of a short-lived convection event lasting only a few days immediately prior to sampling. We would only expect a rectifier signal in the deep firn if the convection persisted for several weeks/months and unfortunately we do not have any isotope data from the deep firn to test whether or not this was the case at South Pole during this sampling campaign. Instead, to demonstrate that convection in the upper firn can affect gas isotope signals in the deep firn, we performed an additional experiment under idealized conditions and included it as an additional figure. We include the new figure and the description from the text below. We also include an analogous figure for the summer rectifier in Section 5.2.4.2.
Finally, the absence of a rectifier in our youngest SPICEcore samples is not necessarily in conflict with the existence of a rectifier in the firn air dataset from South Pole. The youngest sample we analyzed is from 5 kyr BP and our SPICEcore record contains plenty of spatiotemporal variability in ΔT_z (and therefore rectifier strength) on this sort of timescale. Similar to Dome F, recent anthropogenic warming at South Pole may have helped to erase any rectification that existed in the pre-industrial era.

Sturm and Johnson (1991) demonstrated that buoyancy-driven overturning occurs readily in sub-Arctic snow in Alaska. By making hourly observations of the three-dimensional temperature field within the winter snowpack for three years, they were able to observe large horizontal temperature gradients within the snow that were initiated and maintained by columns of rising warm air and sinking cold air. This convection occurred almost continuously throughout two successive winters. There is also ample evidence for air circulation within snow and firn from Antarctica, particularly if vertical cracks allow for fast upward return flow (Giovinetto, 1963; Albert et al., 2004; Fahnestock et al., 2004; Courville et al., 2007; Severinghaus et al., 2010). Unfortunately, direct observations of changes in firn air composition associated with convection are scant since firn air sampling happens almost exclusively in the summer. However, there are published data from a winter firn air sampling campaign at South Pole. In this case, the authors did indeed find that the peak wintertime isotope signal occurred deeper than their firn air model predicted and speculated that this could be due to downward transport of the isotope anomaly by slowly sinking air (Severinghaus et al., 2001). If correct, this would provide confirmation not only of wintertime convection at South Pole, but also that thermal isotope signals can be carried down into the firn by convection without being destroyed by turbulent mixing.
To test their hypothesis, we compare their wintertime firn air measurements from South Pole with values predicted by firn air model runs with and without parameterized Rayleigh-Bénard convection (Figure 7). In the model run without convection, the gases diffuse towards gravitational and thermal equilibrium as they are slowly advected downwards with the densifying firn and occluded in bubbles in the lock-in zone. Because the model is one-dimensional, it is not possible to explicitly simulate a three-dimensional

Rayleigh-Bénard convection cell. Instead, we model just the sinking core of a convection cell, which we parameterise as an 8 cm d^{-1} downward transport of gas between 0 and 20 m. Between 20 and 25 m, the downward transport decays to zero, resulting in mass convergence that would be balanced in the real world by horizontal transport and a return flux of gas to the surface. This approach allows us to approximate how the gas isotopes respond to convection using a one-dimensional model. The model run with downward transport better agrees with the observed wintertime firn air isotope ratios, with the negative wintertime values occurring deeper in the firn than in the model run with no downward advection. The model and the data therefore support our hypothesis that convection can carry seasonal thermal isotope signals down into the firn.

Because isotope data are only available in the top 16 m of the firn, we do not have an observational constraint on the strength of rectification in the deep firn, where ice core signals are recorded. To demonstrate that seasonal convection can affect isotope values in the deep firn, we perform an additional experiment with the firn air model. We simulate the isotope values in the full firn column under idealized South Pole like conditions (110 m thick firn, -51°C annual mean temperature, 7 cm a^{-1} accumulation) and impose a 14 cm d^{-1} downward advection throughout winter (April–September). In the model, the wintertime signal is advected deeper than the summer signal so is not fully cancelled out. This results in a -0.008‰ bias in the annual-mean signal in the deep firn compared to the control run with no downward advection (Figure 8). The bias is of comparable magnitude to the signals in our SPICEcore record, demonstrating that this mechanism could plausibly explain some of the millennial variability we observe.

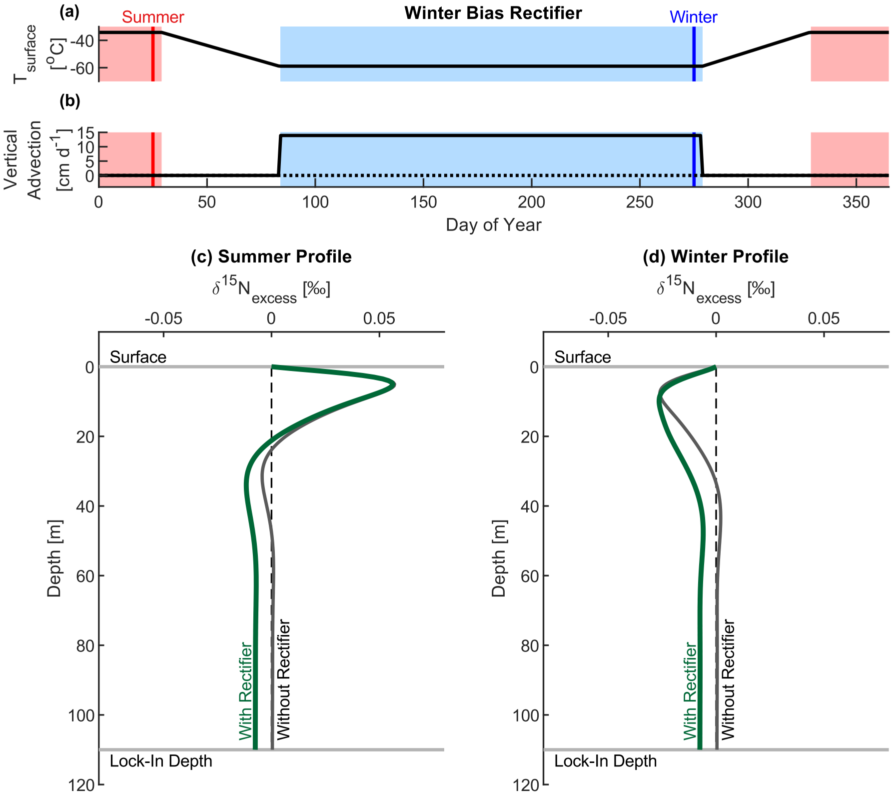

Figure 8. Results of idealized modelling experiment. Panels **(a)** and **(b)** show the temperature and advection forcing applied to the firn air model. The solid lines correspond to the "with rectifier" run and the dotted line in (b) corresponds to the "without rectifier" run with no vertical advection. Panels **(c)** and **(d)** show the vertical profile of $\delta^{15}N_{excess}$ in the firn column at the end of summer and winter respectively. The grey line is the run without advection, the green line is with advection. The days corresponding to the profiles are indicated by the vertical lines in the upper panels.

l-19 and 20: The addition of the Dome F data are confusing and not helpful in this manuscript. It is a different site (much lower temperature). We have many details on the technique for measuring d15N and d40Ar but the data are not show (only firn data and only in the supplement), only DTz from d15N – d40Ar and the DTz from Buizert method but without much explanation on how it is calculated (from which data, with what kind of uncertainties ?). I suggest removing this section which does not add anything to the manuscript. On opposite, figure 7 can be helpful (but need to be shown over the whole firn depth) and is adapted to this study focused on South Pole (see however previous comment).

It seems we did not do a good enough job at explaining the relevance of the Dome F data to the paper. We chose to include the Dome F data as additional evidence that ice core gas records can be affected by rectification. The existence of rectification in multiple ice cores strengthens our argument that these effects cannot be overlooked and that more work is needed to understand where and when rectification is important for ice core gas records. As you point out, Dome F is a very different site to South Pole (colder, higher elevation, dome vs flank site). In our minds, this shows that rectification may be possible over a wide range of site characteristics on the Antarctic plateau.

In order to make this clearer to the reader and to address some of the other concerns you raised, we made the following changes:

- Restructure Section 5.4.2.1, including making the changes to the discussion of Figure 7, as described in our response above and expanding the explanation of the relevance of Dome F to the study
- Expand the Dome Fuji figure to include the Dome F $\delta^{15}N$ and $\delta^{40}Ar$ data, the calculated $\delta^{15}N_{excess}$ and ΔT_z.
- Expand explanation about the Buizert et al. (2021) modelled ΔT_z and add an estimate of the associated error
- Add two new figures showing summer and winter rectification in the deep firn in an idealized firn model (see above).

The paragraphs with the most relevant and significant changes to the text are pasted below, together with the revised version of the Dome Fuji figure.

As further evidence for this type of seasonal rectifier, we also present a previously unpublished ΔT_z record from the Dome Fuji ice core. The core was drilled in 1994-1996 and samples were stored at -50 °C until they were analysed at Scripps Institution of Oceanography in 2007 using a different method to our SPICEcore dataset (Bereiter et al., 2018). Briefly, an ice sample of 800–900 g was melted in an evacuated vessel, and the released air was continuously transferred to a dip tube through a -100°C water trap while stirring the melt water. The air sample was split in two aliquots (Method 1 in Bereiter et al., 2018), one was measured with Thermo Delta-Plus XP for $\delta^{15}N$ and the other was gettered to extract noble gases and then measured with Thermo Finnigan MAT252 for $\delta^{40}Ar$. The isotope data and the reconstructed ΔT_z data are shown in Figure 9, where we compare them to our estimate of the modelled Holocene ΔT_z from Buizert et al. (2021). The model estimate is based on the same firn densification modelling approach described in Sect. 3.3, constrained by $\delta^{15}N$ and empirical Δage datasets described in Buizert et al. (2021). To estimate the uncertainty in the modelled ΔT_z, we re-run the model with different values of the GHF and accumulation rate. We change the GHF by ±10 Wm^{-2} and the accumulation rate by ±10%. The total uncertainty we report is the quadrature sum of the difference between these model runs and the optimal scenario.

Just like the SPICEcore record, the Dome Fuji ΔT_z data show evidence of a wintertime bias due to rectification. The mean of the Holocene ΔT_z data is more negative than both the present-day ΔT_z and the modelled Holocene ΔT_z. Large changes in surface temperature, ice thickness, and GHF can be excluded during the Holocene, so we conclude that the mismatch is most likely due to rectification producing a wintertime bias throughout the Holocene at Dome Fuji. Because katabatic winds are weak at ice domes due to the flat topography, we expect that the wintertime Rayleigh-Bénard rectifier would be particularly effective at this site. This finding strengthens the case for the existence of rectification in Antarctica and demonstrates that rectification can affect gas records at both dome and flank sites and over a wide range of site characteristics (Dome Fuji is 1000 m higher in elevation, 5°C colder, and receives half as much snow accumulation).

Also plotted is the ΔT_z calculated from $\delta^{15}N$ and $\delta^{40}Ar$ measurements on firn air collected at Dome Fuji in 1998, which is -1.2°C (Figure 9, Sect. S3.2). This is more positive than the Holocene ice data and is consistent with the present-day observed firn temperature profile, suggesting no winter rectification is necessary to explain current conditions at Dome F. This could be due to cessation of rectification at some time during the past 2000 years, perhaps in the last century due to anthropogenic warming (the ice surface absorbs downwelling longwave radiation from greenhouse gases very effectively).

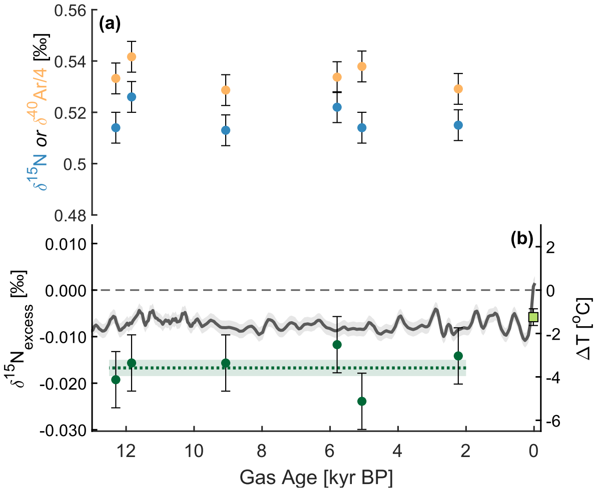

Figure 9. Measurements of (a) δ15N and δ40Ar used to calculate (b) δ15Nexcess and an estimate of ΔTz from the Dome Fuji ice core. The ΔTz data are plotted as dark green circles and compared to a model estimate of past ΔTz at Dome Fuji from Buizert et al. (2021) (grey line and shading). The dashed green line shows the mean of the data and the shading represents one standard error of the mean of the six samples. The light green point shows an estimate of modern ΔTz at Dome Fuji calculated using the method described in Sect. S3.2. The estimate is based on a new firn air dataset from archived samples collected in 1998 (Kawamura et al., 2006) and re-measured at SIO in 2008.

In summary, we propose that low wind speeds over areas of minimal topographic slope cause surface snow temperatures to be colder than on steeper slopes. In winter, this can result in an unstable air density profile in the firn and slow, non-turbulent convection of air to a depth of 10–20 m. This is deep enough to produce a cold, wintertime bias in our ice core records of ΔT_z. In the Dome Fuji ice core, this bias existed throughout the Holocene until at least 2 kyr BP, whereas in SPICEcore, the cold bias is strongest at 20 kyr BP and is co-located with a thicker firn column due to the increased net accumulation of snow associated with slower and/or decelerating winds. Although this hypothesis is somewhat speculative, we believe this mechanism can plausibly explain (i) the most negative values in our record of ΔT_z, (ii) the observed rate of change in ΔT_z, and (iii) the inverse relationship with DCH.

References:

Kobashi, T., Severinghaus, J. P., Brook, E. J., Barnola, J. M., & Grachev, A. M. (2007). Precise timing and characterization of abrupt climate change 8200 years ago from air trapped in polar ice. *Quaternary Science Reviews*, *26*(9-10), 1212-1222.
Kobashi, T., Kawamura, K., Severinghaus, J. P., Barnola, J. M., Nakaegawa, T., Vinther, B. M., ... & Box, J. E. (2011). High variability of Greenland surface temperature over the past 4000 years estimated from trapped air in an ice core. *Geophysical Research Letters*, *38*(21)