Comments to the two reviews and the one short comment by Takuro Kobashi:

We can understand the decision taken by the editor based on the two reviews and the short comment of Takuro Kobashi. At the same time, we would like to mention that we do not agree with most of the comments given by Takuro Kobashi in his short comment as well as Reviewer 2 in this second round, corresponding to Reviewer 1 in the first review round (hereafter referred to as Reviewer 2).

Comment to Reviewer 1:

We are fully aware of the fact that the manuscript is rather difficult to read for a non-specialist. That is why we tried to include as much information as we have gathered during our work on temperature reconstructions using the isotope compositions of nitrogen and argon (and the combined quantity $\delta^{15}N_{excess}$) extracted from ice cores. Obviously, we again failed to transport the complex nature of this reconstruction to the reader. Simply said, and we certainly will take this up in a submission to another journal, it is an issue of signal-to-noise ratio and systematic data uncertainty defining which method should be used for the most robust temperature reconstruction. If we would have access to ideally measured nitrogen and argon isotopes, $\delta^{15}N_{excess}$ would be the best choice. Unfortunately, this is not the case. We hoped to illustrate this issue with Figures 1, 4 and S1 of the manuscript and supplement. Therefore, we have investigated several other possibilities including sole $\delta^{15}N$ or $\delta^{40}Ar$ reconstructions as well as combinations of those.

Comments to Reviewer 2:

Reviewer 2 states (red):

"My main criticisms are as follow: 1. The authors use a rich dataset, with d15N and d40Ar measurements, but make no use of it, and choose to use only one type of data at a time. This is regretable, because they are in fact unable to produce a solution that satisfies all of the constraints (see Fig 4 on the right pannels), even though they could, people (Kobashi, for instance) have done it before"

This is simply wrong; we did investigate the combination of both measurements, i.e. $\delta^{15}N_{excess}$ and made comparisons with Kobashi's (2017) solutions. We must assume that reviewer 2 did not even read the manuscript. In addition, in Kobashi et al. (2017) the remaining misfits for the isotope solutions were not quantified. We recommend studying Figure S1 in the supplement of Kobashi et al. (2017) and judge again.

"2. The authors fail to understand, or explain clearly that, if you fit just d15N, you are basically inferring temperature from firn thickness. If you fit d15N excess, you are inferring temperature from the thermal fractionation in the firn. These are almost independent processes, no wonder they produce different answers."

Here, again we do not know whether Reviewer 2 understands our approach. It is clearly step by step explained in our method paper (Döring and Leuenberger, 2018). We clearly distinguish between firn thickness and temperature influences, the question that we pose in the present manuscript is to which extend are the $\delta^{15}N_{excess}$ usable considering its very low signal-to-noise ratio and systematic offsets toward to too low values (especially in the early and late Holocene). Principally, Reviewer 2 is right when there would not be an issue of uncertainty. However, this is clearly the case as shown for instance by Figures 1, 4 and S1 in the present manuscript.

The reviewer is right by saying: "… no wonder they produce different answers."
Indeed, if you interpret random noise or systematic offsets, it will also result in temperature variations.

3. When inferring temperature from firn thickness (either d15N only or d40Ar only), there are two very important assumptions: 1. firn densification models are perfect, 2. The accumulation scenario is perfect. In this instance of the paper, these two hypotheses have been discussed a bit better, by using two different densification models, and by looking at different accumulation scenarios. When I look at figure 5, I do not conclude that the accumulation scenario does not have any impact, but I al also a bit confused by the fac that the full temperature reconstruction is not shown.

We already were discussing this issue in the first version of the manuscript not only in this second, but indeed we rearranged and incorporated suggestions from you as well as from reviewer 1 on these two issues. We do not state in our manuscript that the accumulation does not have any impact as suggested by reviewer 2! We actually say: "The differences between the accumulation-rates lead to slightly different modelled Δage in the early-Holocene and to a 0.3 K larger cooling for the higher accumulation-rate scenario compared to the two other ones." This sentence refers to Figure 5 of the manuscript.

4. An evaluation of the results, not just in comparison to Buizert and Kobashi, but compared to external validation data, like d18O, borehole temperature reconstructions, other sites etc would be needed to demonstrate the value of this work. Here, we are left hanging with inconsistent results that are not fully interpreted.

It is interesting that Reviewer 2 refers to a comparison of $\delta^{18}O$ of ice. We have to assume that he does not know that Buizert et al., has actually used $\delta^{18}O$ data calibrated to temperature in order to fit the long-term trend of $\delta^{15}N$ to drive their temperature reconstruction. Furthermore, we have compared our temperature reconstructions for $\delta^{15}N$ to the borehole temperature values, see Fig. S5i. A direct comparison between $\delta^{15}N$ derived temperatures and $\delta^{18}O_{ice}$ values have been published in Michael Döring's PhD thesis, see attached Fig. 4.3 below.

To sum up, I don't think that there is enough added value in the work presented here compared to what is already published (the data, the inverse method, temperature reconstructions at the same site) to justify publication. I recommend this article to be rejected.

Reviewer's 2 judgement is based on statements that are simply incorrect when reading our manuscript carefully. It has implications for the temperature reconstructions when using $\delta^{15}N_{excess}$. And, we would like to repeat again that $\delta^{15}N_{excess}$ temperature reconstructions are superior to those using sole $\delta^{15}N$ or $\delta^{40}Ar$ only if the corresponding signal-to-noise ratio is good enough and the causes of the systematic offsets are understood and quantifiable. Otherwise, one had to face significant uncertainties of reconstructed temperature variations but to the point of simply false temperature trends.

Comments to Takuro Kobashi's short comment:

Answer to comment 1:
In Kobashi et al., 2017 it is stated "We circumvented the drifts by allowing slight constant shifts in ΔT by minimizing the difference between the observed and modeled δ15N." This points to an adjustment of reconstructed temperatures based on $\delta^{15}N_{excess}$ to modeled $\delta^{15}N$. If this is not applied, reconstructed temperatures deviate considerably from reasonable temperatures as shown in our manuscript, Fig. 4d. If adjustments are made to $\delta^{15}N_{excess}$ evolution, Fig. 4c hybrid method, then $\delta^{15}N$ plays a very important role again. Also "slight" shifts in ΔT can change the reconstructed temperatures substantially, e.g. when ΔT changes from negative to positive values over an extended time period, the temporal integration will change from a cooling to a warming trend.

Answer to comment 2:
Indeed, Kobashi et al, 2017 has compared temperature reconstructions to several records. But we also did this, including borehole temperature, $\delta^{18}O$-based (Buizert et al. 2018) as well as Kobashi et al., 2017. Therefore, by comparing it directly with Kobashi et al., 2017, we indirectly also compared it to all other records that has been mentioned by Kobashi et al., 2017. In addition, we have attached Fig. 4.3, i.e. a comparison of $\delta^{15}N$ derived temperatures with Gisp2 $\delta^{18}O_{ice}$ values.

Answer to Comment 3:
Yes, the data will be available.

Answer to comment 4:
Yes, we did see Fig. S5i.

Answer to comment 5:
Yes, we did that see Fig. 6. In addition, we add here Fig. 4.3 of Michael Döring's thesis i.e. a comparison of $\delta^{15}N$ derived temperatures with Gisp2 $\delta^{18}O_{ice}$ values as well as a comparison to the Renland/Agassiz $\delta^{18}O_{ice}$-based reconstruction of Vinther et al. 2009.

Answer to comment 6:

The late and early Holocene data show the strongest systematic offsets in $\delta^{15}N_{excess}$ (and ΔT) compared to the rest of the record. This issue becomes visible when inverting the data to temperature (Fig. 4d). We do not say that these systematic negative offsets are caused by gas loss, but we stated that the mechanism which causes these offsets works in the same direction (too low $\delta^{15}N_{excess}/\Delta T$ may be caused by to high $\delta^{40}Ar$). This is also visible in supplement figure S3 in Kobashi et al. 2017 (largest offsets between modelled and measured $\delta^{15}N_{excess}/\Delta T$ in these sections).

Figures from PhD Thesis:

T(δ^{15}N) vs. GISP2 $\delta^{18}O_{ice}$:

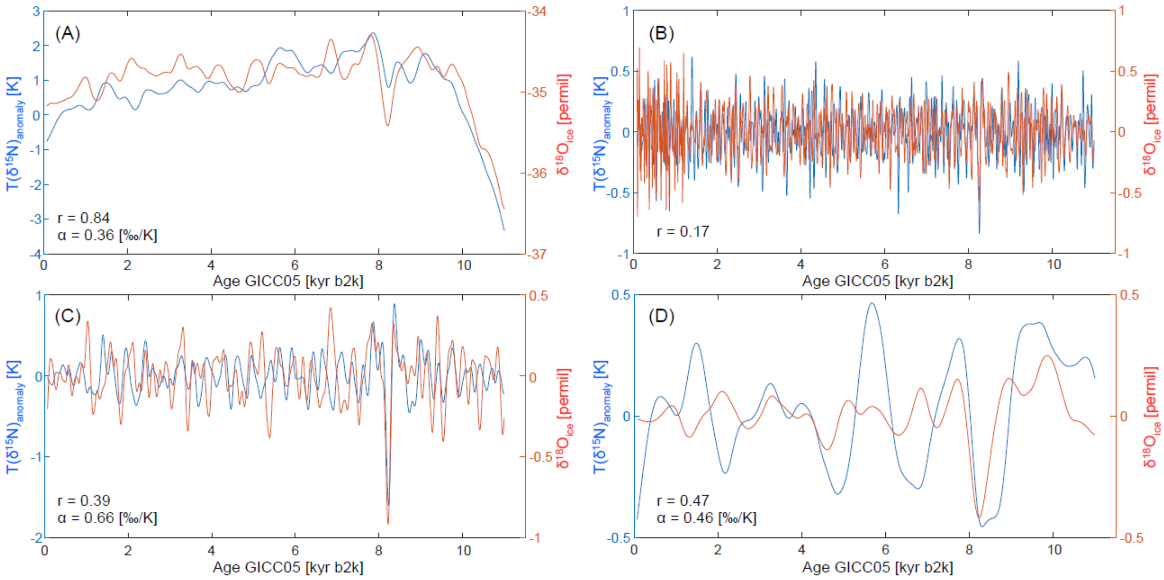
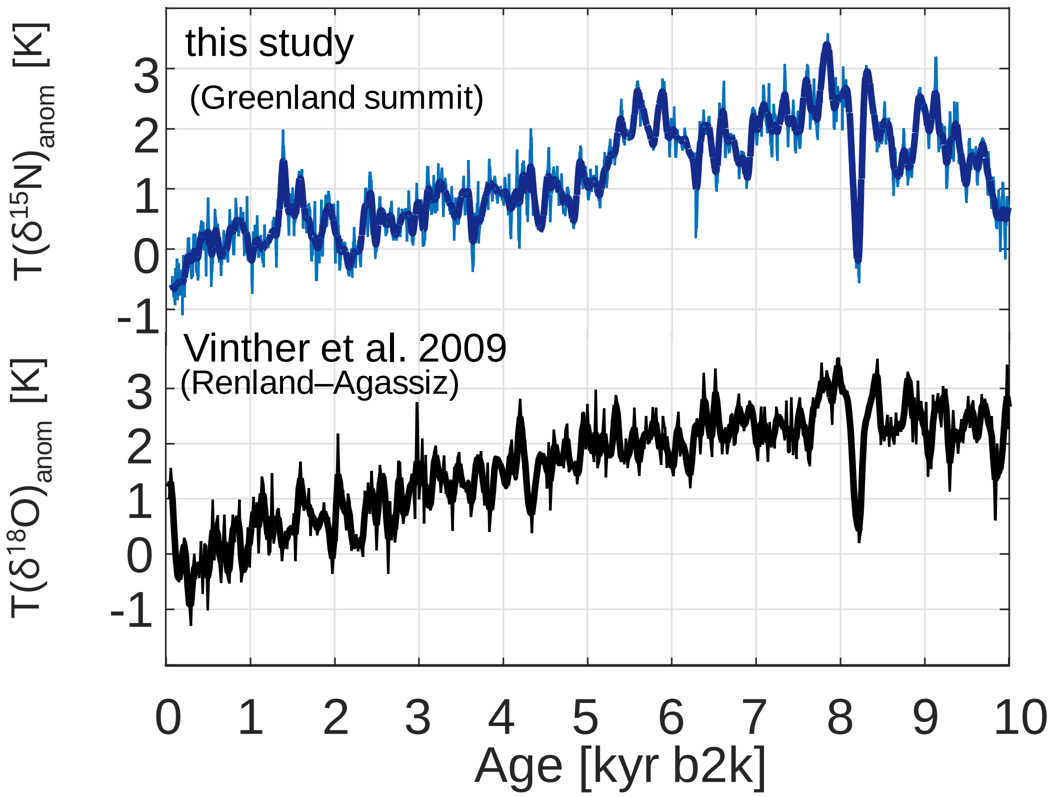


Figure 4.3: Comparison of T(δ^{15}N) anomaly (blue curves) with $\delta^{18}O_{ice}$ (red curves) over the Holocene on different periodic-time bands. (A) Low-pass filtered signals (cut-off-period: 500 yr) display the general trends. (B) Multi-decadal signals (band-pass: 50 yr to 200 yr). (C) Multi-centennial signals (band-pass: 200 yr to 1000 yr). (D) Multi-millennial signals (band-pass: 1 kyr to 4 kyr).

T(δ^{15}N) vs. Vinther et al. 2009:



Comparison of GISP2 T(δ^{15}N) anomaly with stable-water-isotope ($\delta^{18}O_{ice}$) based temperature reconstruction from coastal Greenland sites (*Vinther et al. 2009*).