

Interactive comment on “How well do different tracers constrain the firn diffusivity profile?” by C. M. Trudinger et al.

Anonymous Referee #1

Received and published: 21 September 2012

General comments:

The article presents a new method for the calculation of best estimate firn diffusivity profiles and associated uncertainties. Firn diffusivity is an essential parameter for the modelling of trace gas transport in firn. In my opinion, this study is thus important for scientific issues related to trace gas age distributions and inferring atmospheric trace gas trends from firn and ice core data. I think that the manuscript would be read and used by a larger audience if it was written in a shorter and more goal-oriented way. In its current state, I find the manuscript's main conclusions weak and many contradictory intermediate conclusions in terms of which tracer(s) is (are) best to constrain diffusivity render it difficult to follow. However, I think that the methodology used has a clear potential for providing recommendations for future firn air pumping operations. Important

C7245

issues might easily be addressed through similar model tests than those performed such as which depth resolution should be used? How results are affected by calibration biases? What is the effect of seasonality related firn layering? Such questions are implicitly raised in the model design but not specifically tested in order to be really answered.

I also have some major concerns that are detailed below:

- about the definition of the equifinality criterion used (see comments 2-4, 9 and 27) and other issues related to the model design (17, 26, 31, 32, 35, 40, 57, 64)
- about the important portion of the manuscript devoted to synthetic cases A and B (~8 pages of text, one complex table and 6 figures). The complexity of the design and analysis of these scenarios, and the fact that the ranking of tracer efficiency to constrain diffusivity is obviously depth and real world uncertainty dependent render their usefulness limited in my view. In contrast, DE08-2 results (section 3.4) are presented in only 10 lines and South Pole results (2 boreholes) in 12 lines. A more in depth analysis of model results obtained with "real" data is needed in my view.
- about the lack of clarity and clear conclusions in many parts of the manuscript, which renders it difficult to read.

I make many suggestions below aiming at improving the analysis of the results and reaching clearer conclusions. I hope that they will help the authors to improve this article which I think has the potential to become a very valuable contribution to the improvement of firn modelling.

Specific comments:

- 1 - p17774 l20-23: This sentence implicitly suggests that previous studies over the last two decades paid no attention to uncertainties. It provides a recommendation that I find difficult to implement: all uncertainties (e.g. non modelled processes) are not easily quantified precisely. Please reformulate.

C7246

2 - p17775 l5-13: This sentence describes physical processes that can be modelled (with different degrees of approximation) in the current state of the art. In the context of this equifinality oriented study, I think that potentially important processes that are not or incompletely modelled should be introduced.

3 - p17776 l10-11: The Buizert et al. (2012) study, which details a method of data uncertainty evaluation, states at the bottom of page 4268 (first column) that these uncertainties should not be interpreted in an absolute sense. I think that data uncertainty is not so easily dealt with and deserved more comments. For example, possible calibration biases between firn data and atmospheric scenarios are not easily detected for species with strong concentration gradients in the upper firn. This argument may somewhat oppose to the principle of using mostly/only species with strong concentration gradients in firn to constrain the diffusivity.

4 - p17777 l1-5: A reference provided by the authors (Beven, 2006) presents the estimation of model errors as essential in the equifinality technique. This issue is a major concern for inverse modelling techniques in general and several methods have been developed to estimate such errors (some are discussed by Beven, 2006) and should be introduced here.

5 - p17779 l25: If I understand well, there is at least one exception to the monotonically decreasing diffusivity with depth: the melt-layer case at DE08. Is the monotonically decreasing assumption necessary with the genetic algorithm? If no, what is its advantage?

6 - p17780 l10-13: I am puzzled by the irregular shape of the upper bound diffusivity uncertainty range on Figures 8, 12 and 13. Could it be related to the prior selection of solutions?

7 - p17780 l21 - p17781 l3: How is the melt layer case handled together with the monotonically decreasing diffusivity assumption?

C7247

8 - p17781 l11-13: How genetic concepts such as "breeding", "mutating" and "generation" apply to firn should be introduced.

9 - p17781 l25- 17782 l4: The error range chosen for the definition of equifinality seems somewhat arbitrary to me:

- Although it is very close to the range covered by the 6 models in Buizert et al. (2012) (0.73-0.92 for NEEM-EU), taking into account the new CSIRO model performance, 3 models show results within a very narrow range (0.73-0.74). Buizert et al. underline the fact that the OSU model has limited degrees of liberty to tune the diffusivity, thus the 0.73-0.80 range could be more significant than the 0.73-0.92 range.

- A more pessimistic view of the Buizert et al., 2012 results could come from the fact that no model fits all data points within error bars. A spectacular example is the HFC-134a data point at ~65m depth, which is not matched by any of the models. Therefore I think that the current state of model and data error understanding in the firn modelling community does not allow to clearly define a single robust criterion for equifinality. Could a narrower range of solutions (e.g. 0.73-0.80) be easily re-selected and commented (e.g. compared with Fig.8)?

10 - p17782 l8 - p17783 l8: the method described is not used, thus this one page long discussion could be removed. p17783 l1: why 16 parameters?

11 - p17784 l9-11: Why is the LGGE-GIPSA model diffusivity used to generate the "true concentrations" of a different model? As noted by e.g. Buizert et al. (2012), at the beginning of their Section 4.1, the spread of diffusivity profiles from different models reflects in part differences in model physics, which are compensated for by the diffusivity adjustments. I find the use of the word "true" in the synthetic scenarios A and B context confusing.

12 - p17784 l12-15: I'm not sure to understand how noise is added to the "true concentrations": is it just used to define σ_i or are the "true concentrations" really shifted

C7248

in a trace gas specific way? If yes, is this shift chosen as a single selection of the random variable or are several selections used? I do not see any obvious noise on the "observations" on Figures 2, 3 and 6. It would be interesting to plot the σ_i values used and comment how they compare with the diffusivity-driven envelopes. What motivated the choice of a noise scaled to the RANGE of each tracer? I do not see a connection between this assumption and some characteristic behaviour of experimental/model uncertainties.

13 - p17784 l23-25: is a single random number value tested or are several values used in several simulations to make the test statistically significant?

14 - p17785 l2-3: does the choice of random values affect the estimation of which tracer or subset of tracers is best suited?

15 - p17786 l1-5: How different is the DE08-2 data based closed porosity profile from the modified Goujon et al., 2003 parameterisation used in Buizert et al., 2012? How does it affect the results? What is used for NEEM?

16 - p17787 l6-13:

- l6-12: I guess that Law Dome ice core data are used to constrain the early part of the atmospheric scenarios for other species such as CO₂ and CH₄, the atmospheric scenario estimate and firn diffusivity evaluation are thus somewhat inter-dependent. But if such tracers are also excluded, there will not be much remaining data to constrain the deep firn. Could you comment?

- l12-13: This would be an interesting test to perform on NEEM.

17 - p17788 l27 - p17789 l1 : I do not understand this sentence - gravitational separation is driven by partial pressure gradients (see e.g. Schwander, 1989), why is it affected by the use of Deddy rather than Dmolecular? Trudinger et al., 1997 do not use dispersive mixing and seem to correctly simulate d15N₂ in the lock-in zone (Figures 4-6). Schwander et al., 1993 also simulate a d15N₂ flattening in the lock-in zone without

C7249

Deddy (Fig 4).

18 - Section 3.1.1 : some choices of vocabulary render this section difficult to understand in detail at first reading. I had to take some notes on definitions given in this section, section 2.3.2, Table 2, Table 3 and Figure 2 caption to understand the main results.

- The word "observations" is used for synthetic data with or without added noise

- The definition of the "true" solution is confusing. For trace gas mixing ratios, is it the LGGE-GIPSA model results or the CSIRO model results with the LGGE-GIPSA model diffusivity ? How different are these two solutions?

- I guess that "true concentrations" at p17790 l8 refer to synthetic data without noise rather than experimental data, I'm I right?

- The fairly systematic use of subset numbers in the discussion of which group of tracers is best suited to fit itself and other tracers is confusing for someone who does not yet remember which species is included in which subset.

I also had to read several times Sections 3.1.1, looking at Table 3 in parallel and taking some notes to try to understand what the main conclusions are. In my view, they are:

- diffusivity fairly well constrained with CH₄, d15N₂ and SF₆. As d15N₂ shows little sensitivity to diffusivity (p17790 l11-13), would CH₄ and SF₆ be enough?

- 5 tracers are needed to avoid over-fitting the data (which is quite a lot if the fact that those tracers have reduced uncertainties compared to experimental data is considered) I think that the main advantage of the methodology for building synthetic scenario A is that it allows to properly detect an over-fitting of the data when not enough tracers are used. In my opinion, the discussion of results should be shortened and focused on that point rather than on the ranking of tracers which is directly affected by the wrong amplitude of the uncertainties.

C7250

19 - On p17784 l11, it is said that synthetic A scenario uses a 3.66m well mixed layer, and on p17790 l11, it is said that d15N2 is quite sensitive to the depth of the well mixed layer, could you clarify?

20 - p17791 l26 - p17792 l2: how is the "true" depth of the well mixed layer determined? What is the interplay between this depth and diffusivity? In my view, there is a conflict between the mixed layer concept and the correct simulation of species which show concentration gradients in the near surface firn (CO₂, SF₆, HFC-134a) that is in part hidden by the fact that synthetic A scenario does not use the surface "data" points and the depth resolution used. I doubt that d15N2 data not corrected from the effect of thermal diffusion would be consistent with a well-mixed layer.

21 - Section 3.1.2:

- I have again a difficulty to understand the main conclusions of this section. For example the worst tracer appears to be d15N2 on p17792 l8, 14CO₂ on p17793 l1 whereas p17793 l5 designates 14CO₂ as the best constraint in the lock in zone. The "best tracers" for the upper firn: CH₃CCl₃ and HFC-134a (?) were emitted in the atmosphere only recently and have concentrations close to zero near and below the 14CO₂ peak, which renders them nearly useless in the firn depth range where 14CO₂ is useful. They are thus quite obviously more complementary than better or worse.

- Witrant et al. (2011) performed the same exercise using experimental data and uncertainties rather than synthetic data. Does this change the main conclusions?

- On Fig. 4, results with CH₃CCl₃-only look better than (or in a few cases similar to) results with 10 tracers in the 15-60m depth range except for d15N2. Could you comment?

- On Fig. 4, d15N2-only constrained diffusivity leads to very poor results for all tracers except d15N2 itself, which is also poorly constrained by all other tracers compared to itself. Does it mean that there is a conflict (data inconsistency or model's wrong

C7251

representation) between d15N2 and other datasets?

- I would prefer to see the equivalent of Fig 4 results/analysis for the simulations constrained with "real" experimental data than scenario A synthetic data.

22 - Section 3.1.3:

- The left side of Fig.5 is not discussed and could be removed

- I do not understand why the subset d15N2 + CH₃CCl₃ + 14CO₂ is discussed only in terms of spectral widths and not in terms of PHI_A etc. (it was not presented in Section 3.1.1)

- Besides the above new subset, I do not see what new conclusion is brought by the spectral width analysis. Could this Section be suppressed, the d15N2 + CH₃CCl₃ + 14CO₂ subset presented in Section 3.1.1, and the spectral width analysis shown only for "real" data (Fig 14)?

- I do not understand why much more detailed results are shown and presented for scenario A than for the simulations using "real" experimental data

23 - p17794 l24-25: the "similarity" between synthetic B and "real" data has already been commented in Section 2.3.2. However it misses an element of primary importance: in error (1) sd can be very different from one tracer to another. For example the three species that have non null concentrations in deep firn (CO₂, CH₄ and 14CO₂) have very different uncertainties and this is likely the primary driver for their ranking as efficient tracers.

24 - Section 3.2:

- The reason why values in column "truth" of Table 3 for synthetic scenario B deviates from 1 is not explained. I guess that the main reason is that a single value of the "random" calibration bias was used for each tracer, I'm I right? To which extent deviations from 1 in column "truth" reflect a lack of statistical significance of the results?

C7252

- The analysis of synthetic scenario B results seems rather inconclusive to me. An important point to me in this respect is the fact that 5 tracers were already needed to avoid an over-fitting of the selected tracers in scenario A and there is no subset having more than 5 and less than 10 tracers in scenario B.

- Some differences between scenario A and scenario B have the potential for helping to answer important questions: about the effects of depth resolution of the data, and about firn data versus atmospheric scenario calibration bias but are not tested individually. Thus I think that scenario B related description and analysis could be suppressed without affecting the main conclusions of the manuscript.

25 - p17796 l17 - l23: I would be very interested to read a more detailed analysis of the comparison of the new CSIRO model results with the other models in Buizert et al. (2012). For example, as the diffusivity and 10 tracers concentration results of the six models are provided in the supplementary files of Buizert et al. (2012), it is possible to build the equivalent of column "TenEddy" in Table 4 for the five other models. An interesting aspect of the Buizert et al. (2012) study is to provide a benchmark of model results to which other firn models or new versions of the same models can be compared. In this respect, it would be interesting to provide equivalent result files to those in the Buizert et al. supplement with the new CSIRO model.

26 - Table 4: when comparing columns "Ten" and "TenEddy", I do not understand why $^{14}\text{CO}_2$ is strongly affected by the convective zone representation in the model (it has a nearly flat profile and a low depth resolution in the upper firn). This is not commented in Section 3.3.

27 - Figure 8: I think that the widths of the diffusivity related envelope and the experimental uncertainties should be compared. The much larger width of the diffusivity related envelope in the 20-60m range for CH_3CCl_3 and HFC-134a is striking. In other cases the diffusivity related envelope is much smaller than data uncertainty (e.g. $^{14}\text{CO}_2$ in the 0-60m range) - over-fitted data? This comparison is not easy to view

C7253

when the envelope and uncertainties are small. Some data points remain well outside the diffusivity related envelope - outliers? How dependent are the results on the choice of the equifinality criterion? Are the tracer/depth ranges where diffusivity related envelopes are large the same as those where the results of different models diverge (Buizert et al., 2012)?

28 - p17797 l1-l22: I could not find a strong conclusion in this scenario B versus real data discussion.

29 - p17798 l8-l12: The interest of HCFC-142b and HFC-43-10mee as additional tracers would be enhanced if the enlarged envelopes for CH_3CCl_3 and HFC-134a was due to an inconsistency between the CH_3CCl_3 and HFC-134a datasets (e.g. calibration bias). The consistency of CH_3CCl_3 and HFC-134a datasets could be tested with the "19 representative solutions": is a solution fitting well one dataset bad at fitting the other?

30 - Sections 3.4 and 3.5 DE08-2 and DSSW20K have a much lower sampling resolution than NEEM. An interesting test could be to reduce the resolution of the NEEM data and evaluate how the diffusivity and tracer concentration profiles are affected.

31 - Figure 10 shows model results in firn down to 100 meters depth, whereas on p17786 (l4) it is said that open porosity is zero around 90m. Could you comment?

32 - p17799 l11-l14 and Figure 11:

- I do not see a clear evidence of a convective zone in the d^{15}N_2 data and model results. Could you comment? Is the d^{15}N_2 slope in the diffusive zone deviating from the barometric slope? Is the enlarged envelope on the results in the 30-40 meters range affected by the convective/dispersive mixing effect going throughout it? Section 2.5 mentions simulations with eddy-diffusion and a well mixed layer, how do they compare?

- I am puzzled by the physical meaning of an eddy diffusion term which goes throughout

C7254

the firn. In the upper firn, eddy diffusion is a simplified way of representing fast mixing due e.g. atmospheric pressure variations or wind pumping. DSSW20K and DE08-2 have nearly the same temperature and pressure, thus nearly the same CO₂ diffusion coefficient in free air. Comparing Figures 10 and 11, CO₂ diffusivity in the upper firn is much smaller (twice to a third) at DSSW20K than at DE08-2. What is the physical meaning of a high Deddy contribution combined with a low Deddy+Dmolecular diffusivity? In the deep firn, Deddy is assumed to represent a form of transport in firn layers mostly isolated from the atmosphere (Supplement, Section 5). What is the physical meaning of a Deddy term in the diffusive zone (where gravitational fractionation occurs for d¹⁵N₂)?

33 - p17799 l15-16: deep firn CO₂ is also under-estimated at DE08-2 (and NEEM) but not South Pole, could you comment?

34 - Sections 3.4 and 3.6: Witrant et al., 2011 also modelled DE08-2 and South Pole. Are there significant differences between the results of the two models?

35 - Fig 14:

- The uncertainty envelopes plotted refer to the 19 representative solutions rather than all solutions corresponding to confidence intervals of 68 %, why?

- The plotted envelopes are fairly symmetric at NEEM but very asymmetric at other sites, could you comment?

- Spectral width in deep firn seems to increase regularly with depth at NEEM and DE08-2 but an important slope break (flattening) is visible at DSSW20K, South Pole 1995 and 2001. Is it due to a different model configuration? Does it affect all solutions?

- NEEM and DSSW20K, which have somewhat similar accumulation rates, show similar behaviours approximately down to the lock-in depth (where d¹⁵N₂ stops to increase): a slow increase of the spectral width up to 4 years, then the spectral width increases faster (this is a common behaviour of all sites, which does not surprise me),

C7255

then the increase in spectral width with depth suddenly stops at DSSW20K whereas it continues at NEEM. Is this striking difference well constrained by the data or due to a difference in model configuration (e.g. eddy diffusion, sampling resolution, etc.)?

36 - p17801 l5-l11: A simple test could help being more precise in the comparison between the two South Pole drill sites: using the diffusivities calculated for one site to simulate the other site.

37 - Section 3.7 I am very surprised that the title question: "How well do different tracers constrain the firn diffusivity profile?" is not discussed in this short "Comparison of sites" section. For example, the smallest set of tracers available is CO₂, CH₄, SF₆ and d¹⁵N₂ (for South Pole 2001). These four tracers are available at all other sites. How degraded are the results at other sites if only those four tracers (with "real" data and uncertainties) are used to constrain diffusivity?

38 - Section 3.8: The question raised in this Section: how are the results affected by a possible bias in diffusion coefficients? is very interesting but I have difficulties to understand the meaning of results and their discussion in detail. The main conclusion(s) of this section is(are) not clear to me.

*Synthetic A tests:

- p17802 l3-l8: As the synthetic data were generated using "true" diffusion coefficients and the firn model, allowing the diffusion coefficients to deviate from the "true" value can only deteriorate the results, right?

- Table 3: comparison of columns PHI_A (or PHI_At) Ten and TenDC for CFC-11 and CFC-12. I do not understand why CFC-11 seems quite affected by variations in the diffusion coefficients whereas CFC-12 is much less affected. These species have very similar atmospheric histories and fairly similar diffusion coefficients, CFC-12 concentrations are about twice higher than CFC-11 concentrations thus the synthetic A uncertainties proportional to the species range of concentrations affect more CFC-12. Is

C7256

this due to a lack of statistical significance of the results (p17802 I6-I7: "it has become harder for the GA to locate the true solution")?

- Table 3: for HFC-134a, why is the difference between Ten and TenDC very large for PHI_A but very small in terms of PHI_At?

- The differences between "real" and optimal (model adjusted) diffusion coefficients (e.g. in %) should be provided in Tables 3 and 4.

- Section 3.1.1 (Synthetic A results) concludes that 5 tracers are needed to avoid over-fitting the data. Optimising 7 diffusion coefficients together with the diffusivity profile may thus be an under-constrained problem. Could you comment? Has the calculation been repeated twice (with a different set of random parameters) to check the stability of the results?

- p17802 I15-I18: If I understand well, only the solution with the lowest PHI_A (with diffusion coefficients between 0.9% too low and 2.4% too high) is considered to conclude that diffusion coefficients are well constrained in the synthetic A case. How are the relative diffusion constrained if all "equifinal" solutions are considered?

- Due to their peak shaped atmospheric scenarios, CH₃CCl₃ and 14CO₂ have diffused into and back out of the firn in the past. Are their diffusion coefficients significantly better or less well constrained than those of other species?

* Synthetic B tests: In this more complex case, I have strong doubts about the statistical significance of the results: the way a change in diffusion coefficients can correct for the synthetic B biases on the data likely depends on both the choices of the species dependent systematic errors and the shape of the atmospheric scenario. Synthetic B results for "TenDC" are only briefly commented, thus I think they could be removed from Section 3.8 and Table 3.

* Results with "real" NEEM data. The authors say at p17802 I27-I28 that the reduction in mismatch is highest for SF₆ and CFC-12. In the case of CFC-12, Figure 8 suggests

C7257

to me that most of the mismatch occurs in the upper firn (above the lock-in depth) and could be due to a slight calibration bias between the atmospheric scenario and the firn data. Could you comment? What is the optimal PHI_N if the CFC-12 diffusion coefficient is not adjusted? I am surprised that the CH₄ diffusion coefficient is strongly modified as the CH₄ data are very well fitted, with a narrow envelope, on Figure 8. Does the change in CH₄ diffusion coefficient aim at better fitting CH₄ or better fitting other species together with CH₄ (which species, in which depth range)?

* Results with "real" DE08-2 and DSSW20K data. Compared to NEEM, DE08-2 has less available tracers: only five (just what is needed to avoid over-fitting the data with the synthetic A reduced uncertainties). DSSW20K has eleven tracers but a much lower sampling resolution than at NEEM. Could you comment on how well the solution is constrained? (e.g. based on optimal versus all equifinal solutions). I do not understand why these sites were preferred to South Pole 1995 for the estimation of diffusion coefficients.

39 - p17803 I27 - p17804 I6: I agree with the authors on the fact that the proposed test would be very valuable, especially for NEEM. I would be very interested to see how the results shown on Figure 8 would be modified, and how they compare with the range of model results from Buizert et al., 2012.

40 - Section 3.9 and Supplement Section 5. At this stage I am not able to make an opinion on the fact that a significant conclusion could be reached on dispersion in the lock-in zone for the following reasons:

- I do not understand the model formulation of this process in equation (30) of the Supplement: how dispersive mixing can change the mass of the trace gas species into the mass of air (see also comment 17)?

- Buizert et al. (2012) conclude at the end of their Section 4.1 that models with completely different parameterisations of the lock-in zone reproduce NEEM observations equally well. The first test in Section 3.9 (p17804 I10-15) leads to the conclusion that

C7258

other sites are insensitive to the addition of a Deddy term in the lock-in zone. Thus unless a significant improvement in terms of PHI_N can be reached at NEEM (with real data) by adding eddy dispersion in the lock-in zone (this test is not performed), I do not see how a clear conclusion can be reached.

- p17804 I20-28: I do not understand the aim of the "synthetic A" test - is this to determine if the solution is not under-constrained in a reduced uncertainty frame? What can be concluded for the real data case?

- p17805 I1-I12: Why DSSW20K is chosen for this test if the above results indicate that NEEM is the only site sensitive to Deddy? I guess that the reason why 13CO₂ is not used to constrain diffusivity is that an atmospheric scenario not dependent on firn/ice data (and thus Deddy) cannot be built, then is it possible to constrain Deddy at DSSW20K?

- p17805 I13: It seems quite obvious to me that allowing for dispersion increases equifinality as it adds one parameter to adjust per model depth level. From an equifinality point of view, is there any hope that a dispersion term formulated as a depth-dependent Deddy flux can be adequately quantified?

-p 17805 I14: what clear evidence that dispersion really occurs in firn do we have? From the introduction of supplementary Section 5 and the reference cited, my impression is that:

+ Severinghaus et al. (2010) included eddy-diffusive fluxes in their model as an alternative to no diffusion at all in the lock-in zone, but what all models seem to really need in Buizert et al. (2012) is MOLECULAR diffusion in the lock-in zone.

+ Severinghaus (2012) refers to Buizert et al. (2011) for the CFC-113 based argument, but I did not see it in Buizert et al. (2011 or 2012) and all models (with/without Deddy) have quite similar results for CFC-113 in the lock-in zone.

+ The two models which do not use a Deddy term in Buizert et al. (2012) do not

C7259

fit the 14CO₂ peak very well (one under estimates it, the other over estimates it) but some features of other tracers in the lock-in zone are not fitted by any of the models. The new CSIRO model simulates 14CO₂ within uncertainties without dispersion in the lock-in zone (Figure 8).

+ I do not understand the d15N₂ related argument as several firn models simulate the flattening of d15N₂ without a dispersion term, including Trudinger et al. (1997), see comment 17.

Overall, in my opinion the most interesting question to fully explore from an equifinality point of view is: is there any hope that a dispersion term formulated as a depth-dependent Deddy flux can be adequately quantified? If I understood well the manuscript, the answer seems to be no.

41 - Section 4: Discussion. The manuscript clarity would be much improved if the major elements in this section were provided earlier as conclusions of the different tests/simulations performed.

42 - p17806 I12-I18: This statement should appear much earlier in the manuscript as it has an impact on the understanding of many results. Does not it solve the above question about dispersive mixing? How does it affect the DE08-2 melt layer representation?

43 - p17806 I19-20: please remove "as has always been assumed". As I see it, multi-tracer diffusivity constraint in firn modelling is at a very early stage of development. Moreover, does not the sentence p17807 I23-25 recommend to do what has "always been assumed"?

44 - p17807 I26-I29: Then is convective mixing down to below the lock-in depth really needed to adequately match the data at DSSW20K? In my view, a negative answer to that question would be a good news for firn modelling because I see some physical contradiction between having significant convection throughout the diffusive zone and modelling the effect of gravity in the diffusive zone as at (or close to) barometric

C7260

equilibrium.

45 - p17808 I10-I11: which species could be used that is significantly different from $d_{15}N_2$ and $14CO_2$, and has a well constrained atmospheric scenario?

46 - p17808 I15-I17: Is the uncertainty on $d_{13}CO_2$ really quantifiable in the Dmolecular+Deddy frame?

47 - p17808 I24-I25: To what extent is this due to the definition of the equifinality criterion that is based on the range of results in Buizert et al. (2012)?

48 - p17808 I28 - p17809 I1: I do not understand how these tests really constrain (quantify) isotopic fractionation.

49 - p17809 I10-I13 and Section 5: From the above results, convective and dispersive mixing seem at least very difficult to properly constrain. For example the Deddy(z) values cannot be properly bounded if they are not monotonous (see p17806 I12-18), which is the case if significant dispersive mixing occurs. Is firm modelling just hopeless? Should the overall structure of firm models be completely revised? Is the simple parameterisation with a Deddy term of the complex physical phenomena of convective and dispersive mixing inadequate for firm modelling due to e.g. the non monotonous shape of Deddy(z)? How important are quantifiable convective and dispersive mixing compared to other transport processes in firm?

Comments on the Supplement

50 - p1 I14-15: t' is also used with fixed coordinates (z) at least in eqs. (26), (45), (46), (49) please check.

51 - Equation (10): should not a total derivative be used in the second term?

52 - Equations (17) and (18) could be more simply introduced as a direct consequence of Equation (12), Equation (16) could be omitted.

53 - Figure 1: most of this complex figure is not commented, could you remove the
C7261

uncommented parts or comment more? For example, the effect of compression (I131) could be illustrated by the negative values of db/dz in Fig 1c. I would be much more interested by a plot comparing the DE08 data based and Goujon et al. (2003) equation based porosities.

54 - Equation (25) seems to be a direct consequence of Equations (3) and (19a) in Rommelaere et al. (1997). Section 1.3 could be shortened, why are all equations presented in Eulerian coordinates?

55 - I do not understand Equation (26), which uses t' (moving coordinates) and z (fixed coordinates). How is air conservation determined?

56 - p3 I205-I206: Do you use the Goujon et al. (2003) parameterisation, or its modified version to specify the full close-off depth ($f=0$) from Buizert et al. (2012)?

57 - Equation (30): why does eddy diffusion (or dispersion) change the mass of the tracer into the mass of air? (see also comment 17). Why referring to Severinghaus, pers. com. (2011) rather than Severinghaus and Battle (2006) ? Are the formulations in the CSIRO model and Severinghaus and Battle (2006) equivalent or different?

58 - Extending all terms in Eq. (32), should not the last term in Eq. (33) be zero?

59 - I did not understand how Eq. (40) is obtained from Eq. (39)

60 - I300-I302: then why Eq. (25) determining the air speed as a function of trapping is used?

61 - I302-I306: I do not understand this sentence. Is J_{MRx} (Eq. 42) defined as a diffusive flux which does not include advection and trapping? Does not advection ($u_{\tilde{}}$) affect the mixing ratio time derivative in eq. (43), which includes the gravitational field?

62 - I323-I327: Thus is the upward flux of air finally negligible?

63 - Can you precise how Eq. (48) is obtained from Eq. (47)?

64 - The new CSIRO model (Eqs. 49-51) appears to be Eulerian (z coordinate) rather than Lagrangian (y coordinate), true? How does it affect the representation of a melt layer? Could not the model description be simplified a lot by expressing all equations only in terms of z?

65 - I408-409: and Rommelaere et al. (1997)?

66 - Supplementary Figure 4:

- Several dark bands are visible for some species (several preferred values of the diffusion coefficients?), could you comment?

- Why not show the results obtained with "real data" rather than synthetic B?

Technical corrections:

p17776 l2: "with some kind of automated calibration method" The method by Rommelaere et al. (1997) could be more elegantly introduced as e.g. a conjugate gradient algorithm.

p17777 l25, etc.: "a calibration routine ..." replace with something like "A diffusivity identification (or optimisation) algorithm"?

Remove capital letter in "Synthetic" in titles of Sections 3.1 and 3.2, etc.

p17795 l6 further away?

p17796 lines 11-16 and 24-27: these convective layer formulation related comments could be regrouped.

Acknowledgements: I am very surprised that the NEEM acknowledgement paragraph (see <http://neem.dk/publications/>) is not included.

Table 1: Drill date for the second South Pole borehole is 2001 (not 2011), right?

Supplement I79: unclear notations, not used in the following. Remove?

C7263

Supplement I109: the total porosity s is ...

Supplement I515-516: This definition of the lock-in depth is inappropriate for a model using a monotonous diffusivity (and density) profile.

Interactive comment on Atmos. Chem. Phys. Discuss., 12, 17773, 2012.

C7264