Review: An International Intercomparison of Continuous Flow Analysis (CFA) Systems for High-Resolution Water Isotope Measurements in Ice Cores by Petteni et al.

I. Overview

The manuscript *"An International Intercomparison of Continuous Flow Analysis (CFA) Systems for High-Resolution Water Isotope Measurements in Ice Cores"* by Petteni et al. deals with the comparison of three Continuous Flow Analysis systems developed for measuring the water isotopic composition ($\delta^{18}O$, δD) of ice core samples. The study focuses on comparison tests, providing quantitative evaluations of the resolution and precision achieved by the three systems. Example data from a firn core from Antarctica are used to support some of these tests.

This is a relevant contribution and it fits within the scope of the AMT journal. The manuscript is of good quality in terms of its methods and presentation; however, it lacks clarity and presents some important flaws and misconceptions, especially with respect to the diffusion and deconvolution parts. Therefore, I recommend publication after significant and major revisions are considered.

II. General remarks

I think it is important for the authors to state early in the paper that this is an intercomparison using firn cores. While it is understandable that these systems can be used for ice cores, a significant part of the paper deals with sample diffusion effects, for which the porosity of the core sample is of immense importance. The term "international" should also be reconsidered and replaced with something more appropriate.

The flow of the paper is not entirely smooth, and there are sections from the Methods and Results that could, in principle, be bundled together. For example, Sections 2.6, 3.1, and 3.3 all deal with sample diffusion, signal attenuation, and spectral methods. The titles in their current form are vague and somewhat misleading.

Some of the nomenclature used is atypical or incorrect. The authors use the term "mixing" extensively. This should be replaced with diffusion/dispersion/signal attenuation, as "mixing" describes a very different process governed by different models and mathematical frameworks. The same applies for the term The term "integrated" when referring to the averaging of the data over larger intervals.

An important aspect I struggled with while reading the paper was the sampling resolution of the datasets and the corresponding Nyquist frequency in the plots presenting data in the spectral domain. Please state clearly what your Δx is, and clarify the 1-s sample acquisition time of the CRDS instruments. Typically, Picarro spectrometers export data at non-fixed time intervals. In all plots presenting spectral data, ensure that the Nyquist frequency/period is clearly indicated. Currently, the PSDs of the discrete data, for example, do not extend to 33.3 m^{-1}. What is the explanation for this?

Lastly, some of the central points of the paper regarding noise estimation and sample dispersion are based on calculations and data that are not clearly described and lack mathematical clarity. For example, the authors only present the fits to the impulse responses in Fig. 6 without specifying what exactly the reader is seeing. The Allan variance calculation also lacks explanation, and based on the shape and smoothness of the lines in Figure 4, it is quite clear that it has not been done correctly.

Many of the claims regarding the mathematical treatment and the calculation of the transfer functions involved are incorrect, particularly in Sections 2.5 and 3.3. The manuscript lacks a clearer mathematical foundation to support these claims. The same applies to the comparison between the CFA time series and the discrete data. The evaluation of the results is largely subjective.

III. More specific on the various sections

A. Experimental

The manuscript can benefit from more clarity and depth relating to the desciption of the experimental part. This is very important for a submission to AMT. There is absolutely no information about the three vaporizers, a critical part of the system with respect to precision and achievable resolution. Are they all based on the capillary method? What is the diameter and the length of the capillary. What is the bore diameter of the tee split in the vaporizer? Also information about the specifics of other parts of the system like the filter used. I would appreciate all the diagrams to be moved in the same section in the main text.

Since this paper assesses the performance of the systems based on firn analysis a subsection dedicated to the melters is essential. Capillary effect in the firn can signifficantly affect the dispersion of the sample and the choice of the melter design can have a big impact on these effects. Please provide the drawings of the three different melters. The same should apply for the design of the debubblers.

B. On water concentration and Allan variance

In Section 2.4.1, the authors conflate two distinct effects of water concentration on isotope spectroscopy. The first relates to the choice of water vapor concentration that yields optimal precision. The second, as described in [2], concerns the dependence of the water isotope ratio signal on water concentration, which required a linear correction in that study. The selection of the 15,000–22,000 ppmv range in [2] pertains solely to the linearity of this dependence and the need for correction. However, the instrument examined in [2] is older and fundamentally different from the Picarro variants used in the present study.

The authors choose to show results only from the Venice system. As this is a technical intercomparison study, I believe results should be shown for all three systems and for both $\delta^{18}O$ and δD.

The Allan variance analysis appears problematic. Given the acquisition rate of the Picarro 2130 and 2140 models (2 Hz), the Allan variance curves should exhibit more high-frequency structure than the very smooth lines presented in Fig. 4. It seems clear that something else is being computed. Please consult [2, 4, 7, 8] for relevant plots and formulas. Additionally, a comment on how the authors transition from the non-fixed acquisition rate of the Picarros to a fixed timestep would be appreciated. Are the Picarros "pinged" at a constant interval via external control software, or is the data interpolated post-acquisition? The manuscript mentions that the results in Fig. 4 are based on at least 2-hour injections of UPW. Why do the Allan variance curves stop at approximately 2000 seconds and not extend to at least 3600 seconds, which would be the expected upper limit for $\tau = t_{acq}/2$ [1]?

It would be helpful to see the code used for this calculation, or at least a clear mathematical formulation. The $\delta^{18}O$ and δDtime series used in the Allan variance calculation should also be shown.

For a technical intercomparison paper of this nature, one would expect a deeper analysis of the mechanisms underlying the significantly better precision observed in the Grenoble system. Water concentration level is not a plausible explanation, as it is comparable to that of the LSCE system. To state that "the IGE system exhibits higher precision, attributed to better instrument performance as indicated by Allan deviation" is a rather cyclical argument.

C. Sample diffusion

The diffusion of the sample in CFA systems and the resulting attenuation of the signal power is an important artefact that must be addressed. There are various approaches to this issue, one of which uses spectral methods and transfer functions estimated from impulse responses and/or step functions. The manuscript presents some of these aspects in Sections 2.4.3, 2.6 (whose title should certainly be reconsidered), 3.1, and 3.3. First, I find the term "mixing" misleading, as it technically refers to the blending of different compounds. Therefore, terms like diffusion, dispersion, and signal attenuation should be used throughout the manuscript.

In Section 2.4.3, the manuscript describes how σ_L can be calculated, referring to [6]. However, based on the schematics of the three systems, σ_{sv} is not equivalent to vapor diffusion as defined in [6]. Downstream of the selection valves, there are peristaltic pumps and filters in both the LSCE and Grenoble systems, all contributing to liquid-phase dispersion. The reason why σ_{sv} is equivalent to σ_{vapor} in [6] is that, in that setup, the tubing downstream of the selection valve is minimal and leads directly to a vaporization unit (nebulizer), eliminating the need for a pump. How does this significant detail affect the calculation of sample diffusion in the present study?

Further on, in Section 2.6, the manuscript describes a modelling approach for the spectrum of the CFA data. The approach is problematic, as it assumes that CFA-induced diffusion adds power to the signal, represented by the term P_1. This is not physically possible. Diffusion does not add power—it only removes it.

Another important aspect missing from the analysis is the diffusion induced by discrete sampling itself. A sampling interval of 1.5 cm is roughly equivalent to a Gaussian transfer function with a diffusion length of 0.5 cm [5].

Throughout the manuscript, there is no information provided on the ice core site characteristics like temperature, accumulation and surface density. How do the authors estimate a firn diffusion length of 10–15 cm? Is this ice-equivalent, or does it refer to firn density at the sampled depth?

In Section 3.1, the reader is presented with step functions and impulse responses, but without access to the underlying data or fits. For a technical publication like this, Fig. 6 should incorporate those elements. The information that the authors have used a sequence of firn–ice samples in various combinations is important. Expected differences in diffusion characteristics due to capillary effects and firn porosity should be discussed.

One of the most interesting results of the study—but insufficiently investigated—is why the LSCE system shows more diffusion downstream of the selection valve compared to the segment from the melter to the selection valve (8.6 s vs. 14.4 s, Table 4). The other two systems—and every system I am aware of—show the opposite behavior. Additionally, the LSCE system does not appear to be fundamentally different from the others. This is something the authors should look into.

Regarding Section 3.3 (Spectral analysis), I have several comments. First, it lacks a clear description of the mathematical foundation. The text describes the deconvolution step as an inverse Fourier transform, but it is

not specified what exactly is being transformed. Do the authors construct a restoration filter? Is it optimized for measurement noise as in [3]? The text lacks both mathematical clarity and detail.

There are also misconceptions regarding the influence of the various transfer functions (firn/CFA) on signal attenuation. A transfer function with a diffusion length of 15 cm has a much greater impact (several orders of magnitude) on cycles with periods of 3–20 cm (5–33 m^{-1}) than the CFA transfer function with a diffusion length of 1.5 cm. See plot below. So why do the authors claim that diffusion with a diffusion length of 10-15 cm primarily smooths the climatic signal over periods of 20-50 cm?

The authors claim a significant improvement in the 3–10 cm cycle range due to back diffusion correction, but no data are shown to support this. The data shown in Appendix D indicate the effect is negligible. Which is it?

How is it possible that the measured signals lack cycles in the 10–20 cm range? A quick inspection of both the measured profiles and their power spectral densities reveals significant power in these frequencies. At the same time, a Wiener restoration filter for deconvolving ice core data with diffusion lengths of 13.4 and 16.4 mm is shown in Fig. 2 of [3]. It is clear that both these back diffusion filters—with values very similar to those in the current study—act extensively in this frequency range. Can the authors elaborate?

Clarifying these questions requires presenting the mathematics used—how is the restoration filter constructed, and what does it look like in the frequency domain?
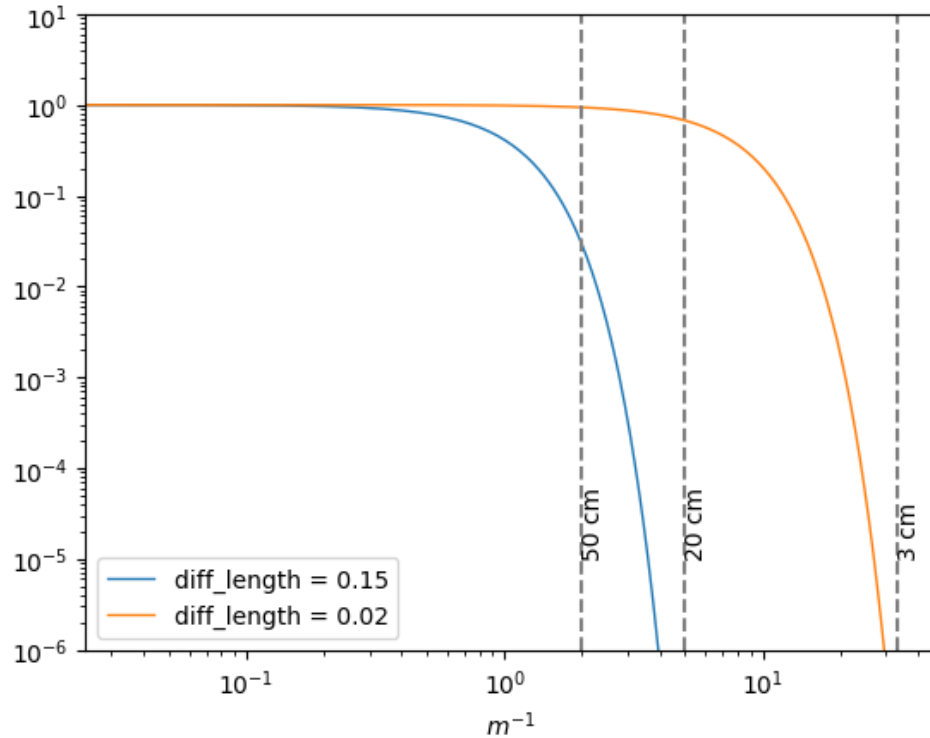
D. Discrete vs Continuous

In the comparison between the produced time series, the terms "statistical difference" and "significant difference" are used. I believe it is important that the authors explain these terms and clarify what objective test they use for statistical significance. A sound normality test for the residuals between all the time series would greatly improve the manuscript. The Shapiro-Wilk and Anderson–Darling tests are some possible choices.

I believe that the manuscript needs extensive work in the review phase addressing these key points, therefore I will not add more minor comments in this review.

References

[1] F. Czerwinski, A. C. Richardson, and L. B. Oddershede. Quantifying noise in optical tweezers by allan variance. *Optics Express*, 17(15):13255–13269–13255–13269, 2009.

[2] V. Gkinis, T. J. Popp, S. J. Johnsen, and T. Blunier. A continuous stream flash evaporator for the calibration of an IR cavity ring-down spectrometer for the isotopic analysis of water. *Isotopes In Environmental and Health Studies*, 46(4):463–475, 2010. doi: https://doi.org/10.1080/10256016.2010.538052.

[3] V. Gkinis, T. J. Popp, T. Blunier, M. Bigler, S. Schupbach, E. Kettner, and S. J. Johnsen. Water isotopic ratios from a continuously melted ice core sample. *Atmos. Meas. Tech.*, 4(11):2531–2542, 2011. doi: https://doi.org/10.5194/amt-4-2531-2011.

[4] V. Gkinis, S. Jackson, N. J. Abram, M. Curran, T. Blunier, M. Halan, H. A. Kjær, A. Moy, K. M. Peensoo, T. Quistgaard, T. R. Vance, and P. Vallelonga. An 1135 year very high-resolution water isotope record of polar precipitation from the Indo-Pacific sector of East Antarctica. *Australian Antarctic Data Centre*, 2024. doi: http://dx.doi.org/doi:10.26179/ygeq-1a95.

[5] Christian Holme, Vasileios Gkinis, and Bo M. Vinther. Molecular diffusion of stable water isotopes in polar firn as a proxy for past temperatures. *Geochim. Cos-*

Fig. 1. Diffusion transfer functions

mochim. Acta, 225:128–145, 2018. doi: https://doi.org/10.1016/j.gca.2018.01.015. URL http://www.sciencedirect.com/science/article/pii/S0016703718300188.

[6] T. R. Jones, J. W. C. White, E. J. Steig, B. H. Vaughn, V. Morris, V. Gkinis, B. R. Markle, and S. W. Schoenemann. Improved methodologies for continuous-flow analysis of stable water isotopes in ice cores. *Atmos. Meas. Tech.*, 10(2):617–632, 2017. doi: https://doi.org/10.5194/amt-10-617-2017.

[7] E. J. Steig, V. Gkinis, A. J. Schauer, S. W. Schoenemann, K. Samek, J. Hoffnagle, K. J. Dennis, and S. M. Tan. Calibrated high-precision [17]O-excess measurements using cavity ring-down spectroscopy with laser-current-tuned cavity resonance. *Atmos. Meas. Tech.*, 7(8):2421–2435, 2014. URL http://www.atmos-meas-tech.net/7/2421/2014/.

[8] P. Werle. Accuracy and precision of laser spectrometers for trace gas sensing in the presence of optical fringes and atmospheric turbulence. *Applied Physics B-lasers and Optics*, 102(2):313–329–313–329, 2011. doi: 10.1007/s00340-010-4165-9.

Fig. 2. Restoration filters