Author's note to editor: We have taken the reviewer's recommendation that Dekabon be presented as the case of a poor performing tubing material for water vapor isotope measurements. The revised manuscript is based on the following major changes.

1. We added a new tubing material, Dekabon, which demonstrates a long memory effect (longer than the commonly used materials) and therefore provides an unacceptable material example. This major change allows us to remove Bev-A-Line XX, a tubing not commonly used in water isotope studies. We have decided to remove Bev-A-Line XX result presentations in the main manuscript because the mixing ratio and isotopic memories are so excessively long that full equilibration takes half a day. Quantifying the memory metrics becomes impractical using the methods designed for the commonly used materials. We would like the reviewers to know that experimental errors are not the cause of excessive memory. It really is that bad. As a service to the research community, we would like to retain a note in the main manuscript that it is not recommended for studies. Evidence, without replicates is provided in the Supplemental.

2. We repeated a subset of experiments (HDPE and Dekabon) with the sources at two different water concentrations. This shows the reviewer how all signals propagate through a "bad" tubing material and shows differences in tubing length when the tubing memory is larger than analyzer residence time. This also gives further insight into whether the direction of the isotopic switch matters, or whether the water mixing ratio change drives memory effects under these conditions. Differences in memory effects with isotopic switch direction have been identified in H_2O varied experiments, but not in H_2O matched experiments.

3. Differences in memory effects based on tubing length and ID have now been identified, following theory. However, we believe our results are too insensitive to make any comments on the relationships between length/ID and surface area relationships in the commonly used tubings.

Author responses are included in red text.

We appreciate the time, suggestions for improvement, and patience of editor Thomas Röckmann and the anonymous referee for this round of revision and throughout this process.

Summary of Comments on amt-2023-56-manuscriptversion3.pdf

Thanks for submitting and improving opon this manuscript, which details a set of experiments aimed to find the differences between the memory effects of various tubing materials. The work is original and valuable for a broad community of research groups performing isotopic composition measurements.

The manuscript has been improved significantly compared to the previous submission. It feels more complete, it reads better, and most importantly uses a clearer/simpler experimental setup. All of my comments (find them in the document below) are minor comments related to minor logical issues or grammatical issues, with the exception of **one major comment related to the inclusion of the Bev-A-Line tubing as a 'not to use' example. The residence times derived for this tubing material are simply unrealistic. I cannot believe it is possible that this tubing type, which is made of PE on the inside** - Note; the same material then the HDPE tubing which is also tested - has a unrecognizebly different attenuation behavior then all other tubing materials tested. If the authors are really confident and comfortable in presenting this as a result, I would expect a strong discussion on the reasons for this outlying behavior. However, in the current manuscript, the discussion of dielectric constants and other physical features merely treat the 5 other tubing types which performed practically identically. Moreover, the supplement suggests that many of the metrics derived by the analyzers are way of for this tubing material compared to all other experiments with other tubing materials. For one**, there hardly seems to be any difference between the two water vapor composition source streams in terms of isotopic compositions**. (when feeding source water through whatever tube for long enough, the water coming out of the tube should become identical in isotopic to the water going in.)

I suspect that something went wrong during the experiment with the Bev-A-line, which causes the unlikely results. A repeat of only this experiment, which confirms or denies the observed behavior would be my strong suggestion before the work is published. When doing one such more test, I would suggest **redoing one of the well behaving plastic tubings to confirm that the system is indeed returning results like expected**. Finally**, if this time around there is Decabon availible, I would also recommend including it as a true 'known not to use' example** to put extra strength to your potential re-observation that Bev- A-Line has totally diverging behavior.

Please find my line by line comments below. Note that the comments are based on an acrobat highlight markup, where some highlights (especially the green ones) do not have a connected comment.

First, we repeated the Bev-A-Line XX experiment with a HDPE control. And we added a change in water mixing ratio to further illustrate the material behavior. Bev-A-Line XX really is that terrible. Unfortunately, the material details are proprietary, and we cannot speculate on the cause of this poor performance. When adding a beaker of warm water to a coil of tubing in a closed container, the water mixing ratio increased, but we are not sure if it is because the tubing is permeable or temperature sensitive. Warming it with a heat gun also increased the water

vapor mixing ratio of the air output, indicating that it may be especially sticky for water and temperature sensitive. We considered leaving the Bev-A-Line XX results in. We are confident in the observations. However, the performance is so bad, it becomes nearly impossible to do the curve fitting in the same way as the other tubing to produce memory metrics and it requires separate figure scales to see the full equilibration time. The error of the memory metrics increases because the shape of the memory is much more stretched out. For this reason, we decided to remove the detailed Bev-A-Line XX analysis. It is plotted in the supplementary material and noted as a poor performing material.

We also want to note that Bev-A-Line XX is not the same as Bev-A-Line IV or V HT , which this reviewer may have experience with. Bev-A-Line IV and V HT indeed have a polyethylene liner, but to our knowledge Bev-A-Line XX has a patented Hytrel® liner. Attempts to find out the exact material of this Hytrel® liner proved fruitless. We have removed Bev-A-Line XX from this version of the manuscript for clarity and conciseness, but we did rerun those experiments. Full source transition was only achieved after about 6 hours.
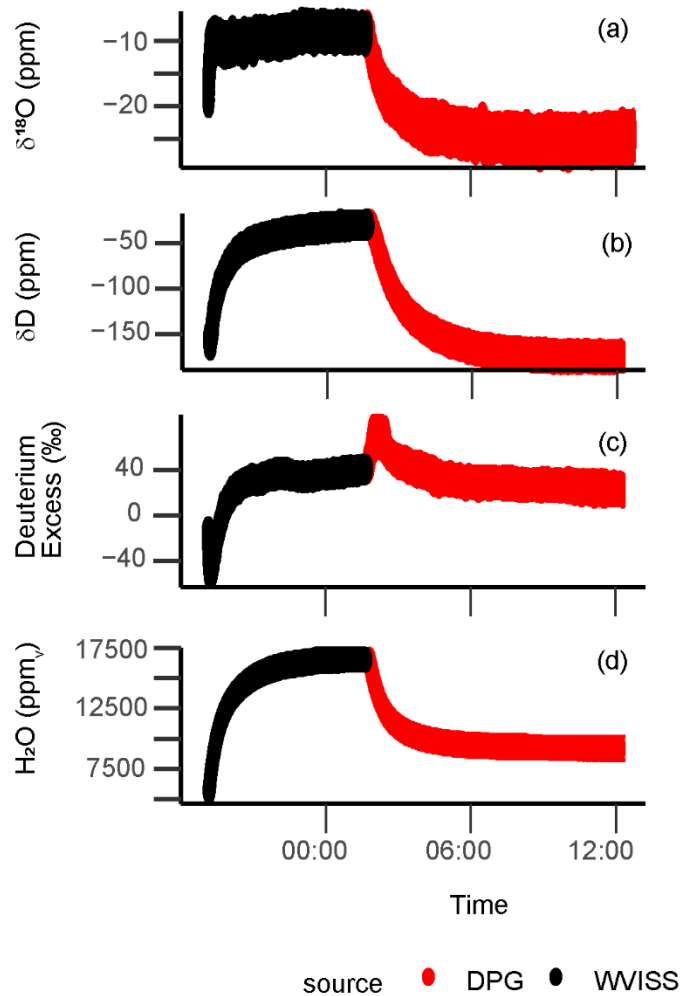
Figure S7. Full sweepout curve of 100 foot (30.48 m) Bev-A-Line XX for $\delta^{18}O$ (a), δD (b), D-excess (c), and H_2O (d) in both depleted-to-enriched (black points) and enriched-to-depleted (red points) directions**.** Bev-A-Line XX takes approximately 6 hours to fully equilibrate in either direction of the switch. The time to equilibration is longer in the enriched-to-depleted direction.

The initial Bev-A-Line XX test from last submission was ran with 1 hour switching which was not enough time for tubing equilibration with the new vapor source. The isotopic values for each source were gathered from a short piece of FEP run immediately prior, reflecting the "true" source values that would have been achieved if the Bev-A-Line XX was allowed to equilibrate in both directions. The values reported in the supplementary table were inaccurate because full equilibrium was not achieved.

Comments in Order (major/red comments bolded)

L. 35 It might be good to specify that the water molecules sticking to the wall are in 'gaseous form', or at least that is what you are testing with the measurement setup.

Added.

L. 38 I get your point, but the effect is not measurement type dependent right?

Correct. Adjusted.

L. 42 new paragraph?

Adjusted.

L. 47 This is seriously implausible.. (referring to Bev-A-Line XX results)

Results of Bev-A-Line XX are correct. Full source transition was only achieved after about 6 hours upon retest. The setup was confirmed with a HDPE control. The reviewer is welcome to test this material themselves. But we would advise against it. See earlier comments.

L. 51 Bev-A-line is also a commonly used plastic tubing material, isn't it?

Bev-A-Line XX is not used in water vapor isotope studies, to our knowledge. Bev-A-Line XX is used in soil gas O_2 and CO_2 studies which is why we had it in the lab and were curious how it would perform for water vapor. Few studies in the emerging soil gas water isotope community have used Bev-A-Line IV, a different material. One could argue that we should have tested Bev-A-Line IV also, but this was outside our budget and time constraints.

L. 76 Note that Bev a line has a polyethylene liner inside. So effectively the surface molecules present i.c.w. HDPE are identical.

This is correct for Bev-A-Line IV and V HT, but as far as we can tell not for Bev-A-Line XX. The XX tubing has a patented Hytrel® liner and attempts to find out what that liner is made of were fruitless.

L. 110 Was this achieved on a synthetic air/vapor source or on ambient air? I think that is interesting for the reader

Synthetic (WVISS produced). Added.

L. 113. Allan variance curves show the entire frequency dependence of the instrument stability. The fact that you choose to display it at '2 seconds' simply means you use the variance (in this case without pre-smoothing the signal). Agree? if so, rephrase. Why you use 2 seconds is also unclear to me here. Why not choose 20 seconds, at the bottom of your 'averaging' curve. That is likely the relevant precision of your analyser.

We have clarified why we have chosen to present 2s allan deviation to quantify measurement uncertainty and refer the reader to Figure S1 to compare analyzer precision. Using a 20 sec running average would smooth away the fast transition times we are aiming to quantify and is counterproductive to this experiment objective. We include it as an estimate of analytical uncertainty.

L. 115 Again, mention if this is a measurement of ambient air or a reference measurement.

WVISS sourced. Added.

L. 120 Inform the reader about the origin of these values. Are these the measurement outputs from the LGR, and post or pre calibrations?

"measured by the LGR TWVIA without calibration" added, Sec. 2.2.1

L. 179 Clear, possibly mention/refer to this when specifying the specs of the standards too.

"measured by the analyzer without calibration" added, Sec. 2.2.1

L. 181 made with or taken by?

Made with. Adjusted.

L. 187 This deserves a less mind braking explanation

Adjusted.

L. 189 idem

Adjusted.

L. 210 I agree partially. The effect smooths the high frequency pertubations, but also introduces a literal lag time, right?

True. In this section we are following the methods for quantifying signal smoothing only. Lags are adjusted for as well in a separate section, but are not the focus of the signal smoothing discussion.

L.263 I immagine that fitting to the transfer function of Bev-a-line was also not self evident?

Correct. Fitting the transfer function was extremely difficult. Since the Bev-A-Line didn't equilibrate in either direction over 1 hour, the transfer/impulse functions were not good fits and should not have been reported in the previous manuscript draft. We have removed Bev-A-Line XX results from the main manuscript.

L.272 Why wouldn't you use the gradient in the transfer fuction itself to extract the derivative in time? Does equation 2 + additions really tell the reader something new, or would a scentence about using the gradient be identical? Please concider!

Without a citation, it's hard to determine what the reviewer is suggesting. There may be some confusion though. The transfer function is fit to the observations to remove high frequency noise which would get amplified by the derivative step. The impulse function is found by taking the first derivative of the transfer function, or d delta/dt, which is the rate of change, which could also be described by the gradient. I don't see that one is superior. Eqn 2 fits the model to the impulse function. The Eqn 2 model fit shown provides additional information in describing the shape of the impulse function following the spirit of the methods set out in Steen-Larson et al, 2014; Jones et al, 2017; and Kahle et al., 2018. So to be clear, the impulse function is the derivative (or gradient) of the transfer function, but the shape of that impulse function is examined by fitting the model (Eqn 2) and reporting the parameters that determine the shape of the impulse function.

L. 320 Could a discrepancy between theoretical and real tube roughness cause this slight difference? Note that Iassume that roughness was taken into account for calculating the flowrates here!

Flow rates were measured directly just upstream of the analyzer and not calculated. But roughness may influence the signal transition by creating more mixing at the transition front.

L350. Again, this is extremely implausible. (referring to Bev-A-Line XX results)

See L. 378 response

   L. 365 Conclusion / discussion? (ALM: Sentence in Sect. 3.2 is "Given differences in D-excess values between sources, we caution overinterpreting the maximum D-excess anomalies between experiments, as evidenced by the different starting points in Fig. 3e.")

Moved to Sect. 4.3

L. 378 You indicate that you scale each 0 and 1 to the beginning and ending values of the transition curves. This is clear for all tubing types, but for bev-a-line it is clear that neighter 0 or 1 is ever reached. How can this be if you rescale?

More importantly, appendix table S1 indicates that the averaged isotopic compositions before and after are simply uncomparable to all other experiments. It cannot be possible that this PE tube is so porous that you cannot measure the isotopic composition of a constant source through it after equilibrating for hours.

Full source transition was only achieved after about 6 hours.

See earlier comments.

L. 406 repeated use of word 'different'. Also, the message of the sentence is unclear.

Fixed. "Each memory metric calculated provides a different ranking of tubing material based on slight numerical differences in metric values, and all tubings appear operationally similar with the exception of Dekabon (Table S2)."

L. 407 Strange use of However, after a scentence that already has an 'however' structure!

Fixed. "Some common patterns in these rankings do emerge."

L. 431 Where are the D-excess curves for BEV-A-Line?

Good catch- I forgot to add those. This has been fixed for Dekabon.

L. 432 This is not possible. All the D-excess t3 trashhold values occur after t=5 seconds. Possibly your algorithm finds a zero crossing instead of a zero crossing with a negative slope. (ALM: referencing "Finally, residence time adjusted t3‰ values for D-excess range from 0–93 seconds, while the largest t3‰ uncertainty value was 536 seconds.")

We agree that the 3 per mil threshold is reached after 5s in Figures 3, 4, and S3. The 0 seconds for D-excess threshold value was residence time adjusted unheated short thick-walled FEP (WVISS direction). In the location adjusted Figure 4s, the D-excess anomaly never exceeds 3 per mil. In this tubing and transition direction, the memory effect is so short there is basically no significant D-excess anomaly, leading to us to report a D-excess t3 value of 0 in the manuscript. However, we see the confusion, and have listed this value as N/A now in the memory metric table, and this sentence no longer appears in the manuscript.

L. 435 make a proper scentence.

Sentence removed.

L. 443 I suggest putting this subchapter under the discussion, as its not your findings that are discussed but the findings of other researchers.

Fixed. Now Sect. 4.1. We do wish to point out that we are amenable to removing this section if the editor agrees it does not add to the manuscript.

L. 451 If this is the reason we observe the difference, why does the attenuation time in no way scale with the dielectric constant?

It is possible the dielectric constant is not a good characterization of wall effects for water isotopes. See Sect 4.1 for further discussion into this topic. We do wish to point out that we are amenable to removing this section if the editor agrees it does not add to the manuscript. We were searching for predictive capability to screen tubing materials, but unfortunately, did not find that these metrics were useful.

L. 453 When evaluating material types like you do here, why not mention Bev-A-line and try to explain why it behaves so strangely? It is effectively PE, so the same dielectric properties HDPE are likely found.. This is a major missing point.

See previous comments for Bev-A-Line. We do not attempt to speculate on the reason material types are different- we had a hypothesis on what properties may play a role in isotopic effects but they do not appear to play a significant role based on our results.

L. 457 include Bev-A-line! (ALM: now Dekabon)

We cannot include Dekabon (or previously Bev-A-Line XX) due to the undisclosed patented inner liner material.

L. 477 Why are only slight differences predicted based on material? Copper is vastly diffident that the plastics used in terms of properties. If theres properties were related to the ultimate attenuation you find, this shouls be seen right?

Changed to "may predict differences". The memory effects of these tubings (with the exception of Dekabon) are very small relative to the properties of the analyzer. The properties we reference may impart non-fractionating effects, or the memory effects may be too small for these properties to have a visible impact. But we bring up % water absorption and relative permittivity as two of the material properties that seem like they would have an impact. Further research is needed into how material properties of tubing affect water isotope memory. We do wish to point out that we are amenable to removing this section if the editor agrees it does not add to the manuscript.

L. 479 I do not agree this can be concluded. Material properties dit differ, and 2x 20s is significantly more then 4s...

While material properties do differ, we are not attributing these differences to any operational differences in tubing performance. There are few consistent differences between tubing

L. 493 but not the majoritie of the difference right? could tube roughness play a role here?

L. 499 What are these wall effects governed by if not temperature variations, temperature, or material properties?

L. 520 Strange thing here is that decreased ID increases the relative surface area but decreases the absolute surface area.

L. 525 The metrics are not slow. decreased memory times?

L. 525 repeat of also

L. 526 Time is not fast, but short of long.. please improve such logical errors to improve the overall readability!

L. 533 Please rewrite every point where Bev-A-line is excuded from the analysis / conclusions without proper reasoning. How come that Bev-A-Line very strongly does not follow this expected behaviour?

L. 542 0.7x is not greater then. I suggest changing the wording.

L. 577 I here mis the obseravtion that the center point of a pulse will be timedelayed w.r.t., for example, the molefraction variation of H2O itself. So this attenuation imposes both a lag time as a smoothing effect.

We now show in our results that isotopic values transition slower than the H_2O signal (Section 3.4), and have adjusted the wording to include lag time and smoothing.

L. 629 Spectral? with laser spectrometers you mean? Mass spec is also spectal.

Changed to "laser-based spectral isotopic analysis"

L. 635While this is certainly often the case, it is not nececarally true.

Added "often" to sentence. Now reads "Liquid water analysis is one example of a case where air flow rates and temperatures of transfer lines are often fixed by the instrument design."

L. 647 What does the XX specification indicate? may be good to mention somewhere!

This is the manufacturer's name and is not elaborated on. As far as we can tell, the XX specification is just the numerical value added to the end of the Bev-A-Line name to distinguish it from other Bev-A-Line tubing lines, including Bev-A-Line IV and Bev-A-Line V HT.

L. 652 I think it is important to mention that your RH is in any case not above 50%, even without heating the inlet. As far as I understand, the community uses heated inlets when RH~95 may occur without heating, which starts to form a risk for liquid water films.

Added "At this relative humidity of ~33 %" and "Heating to avoid condensation does not seem to negatively impact the isotopic measurements."

L. 655 In which direction? does this allow you to draw conclusions on the relative surface area (in m per m2 flow area) vs absolute surface area importances?

Increasing tubing ID and length increased the threshold metrics. However, we have concluded that the experiments are not sensitive enough to confirm linear or non-linear behavior with surface area. Dependence on length and ID would suggest memory is a function of total surface area or number of exchange sites.

L. 664. This suggests some randomness that I do not believe is supported by your data or the exchange principles you detail.

This sentence has been cut.

References:

Steen-Larsen, H. C., Sveinbjörnsdottir, A. E., Peters, A. J., Masson-Delmotte, V., Guishard, M. P., Hsiao, G., Jouzel, J., Noone, D., Warren, J. K., and White, J. W. C.: Climatic controls on water vapor deuterium excess in the marine boundary layer of the North Atlantic based on 500 days of in situ, continuous measurements, Atmospheric Chemistry and Physics, 14, 7741–7756, https://doi.org/10.5194/acp-14-7741-2014, 2014.

Jones, T. R., White, J. W. C., Steig, E. J., Vaughn, B. H., Morris, V., Gkinis, V., Markle, B. R., and Schoenemann, S. W.: Improved methodologies for continuous-flow analysis of stable water isotopes in ice cores, Atmospheric Measurement Techniques, 10, 617–632, https://doi.org/10.5194/amt-10-617-2017, 2017.

Kahle, E. C., Holme, C., Jones, T. R., Gkinis, V., and Steig, E. J.: A Generalized Approach to Estimating Diffusion Length of Stable Water Isotopes From Ice-Core Data, Journal of Geophysical Research: Earth Surface, 123, 2377–2391, https://doi.org/10.1029/2018JF004764, 2018.