We thank the reviewers for very useful comments. We provide below the answers to the different comments and revised the manuscript accordingly. Note that part of the answers as well as the need for a more detailed explanation of the theory associated with the vaporization of the drop is provided in a companion paper which is now available online (Kerstel, AMTD, 2020). We also added one co-author, Morgane Farradèche, who realized numerous of the additional tests requested by the reviewers.

Reviewer 1

Technical comments: 1- This study represents an extension of the work shown in Landsberg (2014). However, in his study a strong influence of the lead screw rotation was identified as one of the main source of noise in the water vapor signal at low humidity. Is the choice of the pump a critical point for the development of the instrument? There are other critical points for designing/replicating this device that must be taken into account (e.g. how important is the choice mass flow and pressure controller? The authors should discuss this point e.g. in section 2.2.

Indeed the choice of pump is important. First of all because it needs to provide an exceedingly small water flux, and secondly because it needs to do so with high stability. It is therefore noted that the oscillations that are in-phase with the lead-screw rotation were seen by Landsberg (2014) in an early prototype and were eliminated by better engineering, and in particular by the careful elimination of air-bubbles in the liquid water supply lines and syringe. This aspect is also addressed in our companion paper (Kerstel, 2020).

The precision of the pressure controller is not critical. Its purpose is to provide a steady inlet pressure to the optical spectrometer, also during syringe switching. Ideally the spectrometer is not sensitive to the inlet pressure.

The precision of the flow controller on the other hand does directly influence the precision of the volume mixing ratio (humidity level) produced by the instrument. Its precision ranges from about 0.5% at the highest flow setting, to almost 50% at the lowest air flow setting. In practice the device is used with an intermediate air flow at which the precision of the controller is specified to be 1%. We observe, however, that in practice its short-term precision is much better. In any event, a precision of the order of 1% is normally comparable to the precision of the measurement of the humidity level with the optical spectrometer.

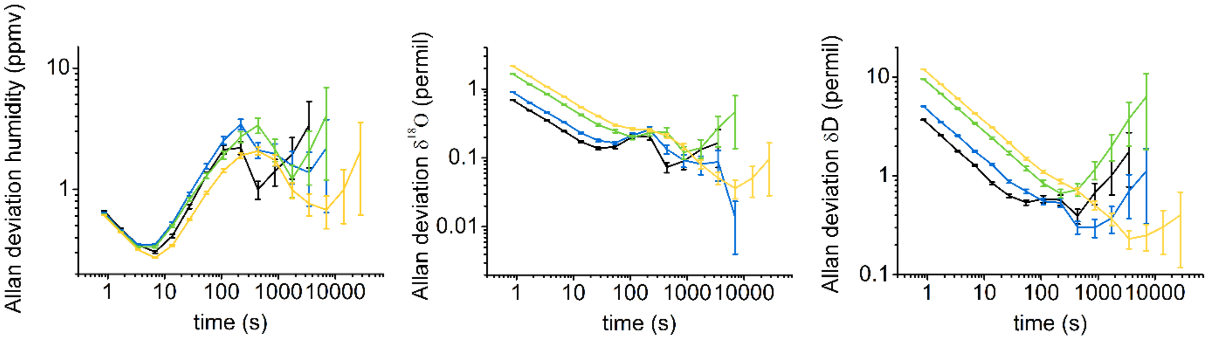
The text has been modified accordingly to provide quantitative information on these issues and cite the companion paper.

Section 2.2

"The spectrometer is not sensitive to the inlet pressure, the precision of the pressure controller is not an essential aspect. On the contrary, the precision of the flow controller is key for the precision of the humidity level produced by the instrument: it is of 1% for the air flow which is comparable to the precision of the measurement of the humidity level with the optical spectrometer."

2- Since large stability of water vapor flux is expected by such device, one would expect an analysis of the stability of water vapor signal, at least in terms of the mixing ratio. An analysis of stability could be the identification/absence of trend during different humidity steps or the analysis of mixing ratio standard deviation as a function of different instrumental configuration (e.g. dry air flux, syringe speed etc). A stability analysis would provide additional evidence of the robustness/reliability of the instrument.

This issue was addressed through two different tests. First, we now present Allan variance obtained over 4 hours at different humidity levels by varying the syringe speed (Figure 3 of the new manuscript)



(a) (b) (c)

Figure 3: Allan variance over 4 hours for different humidity levels (black 1080 ppmv; blue 770 ppmv; green 400 ppmv; yellow 320 ppmv) for humidity (a), $\delta^{18}O$ (b) and δD (c).

Second, we also tested the stability of the system by varying the flow of dry air from 200 to 400 sccm (Table S2 of the manuscript). The reported standard deviation is over a period of 10 minutes (last 10 minutes of the plateau) but the humidity and $\delta^{18}O$ values were stable over one hour (the experiment was stopped after one hour of stable humidity). We notice that the generated $\delta^{18}O$ values does not depend on the dry air flow nor on the infusion rate.

Air flow (sccm)	Infusion rate (µL/min)	Humidity (ppmv)	stdev humidity over 10 minutes (ppmv)	$\delta^{18}O$ (‰)	stdev $\delta^{18}O$ over 10 minutes (‰)
200	0.07	808	1	-7.88	0.89
300	0.11	851	2	-7.73	0.85
400	0.14	818	2	-7.95	0.90
200	0.03	374	1	-8.45	1.92
300	0.05	411	2	-9.16	1.64
400	0.07	415	3	-9.05	1.59

Table S2: Evolution and stability of humidity and $\delta^{18}O$ (same water used for the different tests) for different syringe infusion rates and dry air flows.

Other minor comments

L136 Period.

Done

L147-148 What is the reason for using fractionation factors of Cappa et al. (2003)?

In this paper we are interested in the steady-state behavior of the device. In this case, the isotopic composition of the incoming liquid water is by mass balance considerations equal to the isotopic composition of the evaporated water flux. The fractionation factor between the liquid and vapor phase is thus of no importance for the work presented here. The previously mentioned companion paper, however, deals with the dynamic behavior of the humidity generator and discusses the fractionation factors in detail. There we use estimations of the fractionation factors that are based on the studies by Cappa et al. (2003) and Luz et al. (2009). The work concludes that the effective fractionation factors are largely determined by their equilibrium counter parts, already determined with high precision by Horita and Wesolowski (1994), combined with the ratio of diffusivities and a parameter that describes the flow regime (from laminar to slightly turbulent). The diffusivities were determined by Merlivat (1978), Cappa et al. (2003) and more recently by Luz et al. (2009). We used the values by Cappa et al. (2003), but the difference is minimal for our purpose. Cappa et al. (2003) predicts only slightly lower values of the effective fractionation factor in the laminar limit compared to other estimates.

As the fractionation factors are not necessary for this steady-state study, we will remove the reference to Cappa in this study. The dynamical aspects are dealt in the companion paper.

Figure 2. Check part labels (A/B) because different names are used in text, in table 1 (F1, G1 etc..) and in Figure 3.

The scheme has been fully changed, see answer to reviewer 2 and new Figure 2 and Table S1.

L245 In relationship with my technical comment #2: what stable means? No trend observed in mixing ratio? Low RSD?

Indeed, stable means both no trend in the mixing ratio (< 20 ppmv/hour) and low RSD (< 10 ppmv over 10 minutes). This is explained in the revised manuscript.

“The LHLG is able to generate stable levels of humidity (drift lower than 20 ppmv over one hour and 1σ below 10 ppmv over 10 minutes)”

Figure 4. It is not clear how long it takes the signal to stabilize and whether this stabilization period is related to the humidity level. From Fig.4 it seems so, because injection steps are characterised by different lengths.

Actually, it can be variable between a few minutes only to 30 minutes in worst cases. In best case, we have stabilization in a few minutes but it may take more time for undetermined reason. This is the reason why, in routine mode (i.e. when the user is not checking visually the stability of the plateau), we run 50-minute long plateau. This is explained in the revised text.

“In the routine mode (Figure 4), we perform plateaus of 30 to 50 minutes (50 minutes when the instrument is unattended since the time to reach the plateau varies between a few minutes to 30 minutes).”

Section 3.2 Maybe, two words here (or a simple scatter plot, boxplot) about a possible relationship between precision and humidity level could help the reader understand the stability of the system.

There is actually no relationship between precision of humidity measurement and humidity level as seen in figure 3a (new in this manuscript).

Section 3.3, Figure 5 and Section 4: humidity correction is important. I would like to point out to the authors that a correction based on the difference between observed and reference values for a single standard might not be enough, as recently highlighted out in Weng et al. 2020. Under the perspective of the influence of both humidity and isotopic composition, the correction function should represent a plane and not a line (as one could expect from Figure 4). Moreover, from Figure 4 it is not clear if the scatterplots reports the difference between a single standard (obs-ref) or for different standards (with different isotopic composition) because the plural "standards" word was used. Maybe the authors could highlight that under low humidity and low isotopic variability, the humidity response of the instrument can be determined by analysing one (or two) standard(s).

Indeed, we paid attention to this calibration issue and ran several standards. For standards with low δ^{18}O values (-30 to -55 ‰) used for calibrating our instrument in Antarctica, the relationship between δ^{18}O (δD) and humidity is the same which makes the calibration of the laser analyzer easier in this region. On the contrary, when we use a less depleted standard (δ^{18}O =-9 ‰, δD = -43.6 ‰), the relationship between δD and humidity is not the same as displayed on the new figure 5.

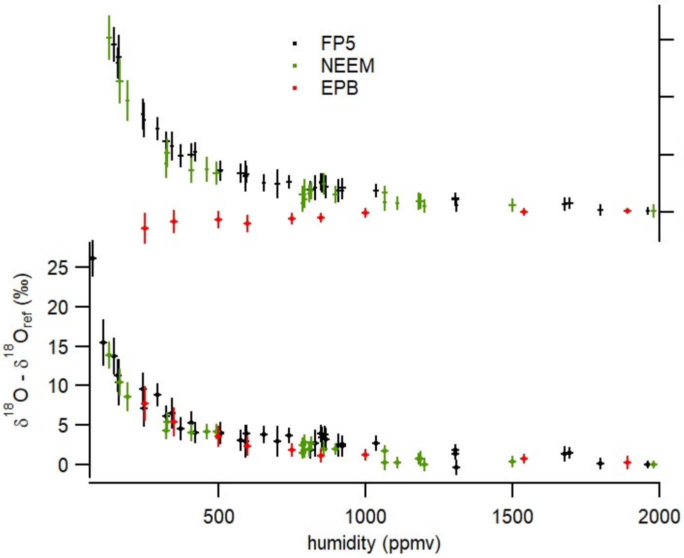


Figure 5*: Influence of humidity on the isotopic composition (δ^{18}O and δD) of the vapor obtained with the LHLG with 3 water lab-standards. The error bars are calculated as the standard deviation (1σ) over the generated values by the L2130-i instrument during 10 minutes at 1 second resolution (i.e. without any pre-averaging of the raw dataseries). The $\delta^{18}O_{ref}$ and δD_{ref} are the values of the injected water standards at 2,000 ppmv.*

Reviewer 2:

Remarks

2.1 Evaporation Model

The theoretical foundation of the vapor generator presented here is described in section 2.1 of the manuscript. It is based on the assumption that a constant flow of water ending on a syringe needle tip combined with a finely controlled flux of dry air will produce an isotopically stable stream of vapor via a zero fractionation process. A prerequisite for this, is that the size of the droplet remains stable throughout the experiment. It is still a question to me how a droplet that sustains its size by constantly loosing mass on its surface (regained by the incoming liquid water flow) via evaporation yields a vapor stream that has the same isotopic composition as the liquid. This is a typical Rayleigh evaporation likely with a strong kinetic component due to the very low humidity of the carrier gas and the quick -but incomplete- removal of water molecules from the droplet. Even though the treatment of the model has been presented in previous works, this is so central to this system that at least an appendix with more information is needed.

We thank the reviewer for raising the question as it is indeed at first sight not immediately obvious that the we can be certain that no fractionation factor is involved in the determination of the isotopic composition of the output vapor. We have therefore decided to submit a companion paper that describes and validates a model of the dynamic behavior of the humidity generator (Kerstel, AMTD, 2020). This manuscript is still in the reviewing process. In short, the mass balance model convincingly demonstrates that the isotopic composition of the generated vapor is identical to that of the incoming liquid, in steady-state. It also shows that in steady-state the droplet volume is constant and that the enrichment occurring at the surface of the droplet due to evaporation of its surface water is compensated by an inward diffusion of heavy isotopologues. This diffusion process, however, is frozen in place by the continued renewal of water arriving in the droplet through the syringe needle. This is important, as otherwise the bulk syringe water would become enriched. Measurements performed on the syringe water before and after several hours of operation show that this is indeed not the case, as predicted by our model.

We will add a brief note to this issue in the text and refer to the companion paper for more details and insight. This also addresses the reviewer’s request for an appendix, which we have thus decided to replace by the companion paper.

“By solving numerically the differential equation (2), it is possible to faithfully simulate the behavior of the device under changing conditions (Kerstel, 2020). This numerical approach validates the theoretical explanation of the undersaturated evaporation of the droplet. Importantly, it is noted that in steady-state, the isotopic composition of the generated humid air is identical to that of the injected water stream, and therefore does not depend on the infusion rate, nor on the specific humidity.”

2.1.1 Some points to consider

- How is the size of the droplet controlled when there is no camera or some other monitoring tool inside the evaporation chamber?

The humidity level generated by the device is controlled in feed forward by the ratio of water to air flow. The droplet size simply follows. There is thus no need to actually measure the droplet size. This said, an early prototype contained a camera that allowed us to produce the photograph in Figure 1.

- The manuscript mentions in Line147 that the fractionation factor by Cappa et al (2003) are used. But why if there is no fractionation?

We thank the reviewer for pointing out that the mention of Cappa, or other fractionation factors, is superfluous here.

We eliminated the mention of Cappa et al. in the revised version of the manuscript, and instead refer to the companion paper for a discussion of the dynamic behavior of the device.

- The manuscript claims that under steady-state the isotopic composition of the generated vapour does not depend on the injected flux of water nor the specific humidity. Since this is a technical paper I would appreciate a simple experiment where the flux of the carrier gas is ramped up and down allowed to reach steady state. I have a very hard time seeing how this experiment will produce a vapor stream of the same isotopic composition equal to the _$_{18}$O of the liquid water. Similarly if the specific humidity of the carrier gas is changed one ends up with a simple mixing experiment (Mook , 2000) where the resulting vapour isotopic composition naturally will be different.

We followed the suggestion of the reviewer and run experiment with different dry air flows. Because the air is dry, it is not expected to modify the isotopic composition of the produced water vapor if the humidity is the same. The results are shown in the new manuscript on Table S2 displayed above as answer to reviewer 1.

- Throughout the whole manuscript, I have not seen a note on the temperature of the evaporation. This is a critical parameter affecting the efficiency of the evaporation, the saturation vapour pressure of the mixture and the fractionation factors in the (Cappa et al , 2003) parameterization. Since there is no active control and regulation of the temperature, all these parameters will vary.

We thank the reviewer for bringing up this issue, as the temperature is indeed a factor to consider for several reasons. Temperature variations affect the stability of the device by their effect on the stability of the flow regulator, which in turn affects the stability of the humidity level. Our flow regulator having a temperature sensitivity of 0.05 %/°C, this factor is deemed not critical. Temperature also intervenes through its effect on fractionation and the evaporation rate, as mentioned by the reviewer. Therefore, it can force a departure from steady-state operation, which in turn would drive isotope variations. For these reasons, the device is carefully insulated, leading to a temperature of the evaporation chamber of about 20°C varying within 1°C over 24h.

In the new text, we precise:
“Temperature intervenes through its effect on fractionation and the evaporation rate (apart from a negligible effect on the flow controller stability), which could lead to a departure from steady-state operation. For these reasons, the temperature of the evaporation chambers was maintained at 20 °C (within 1 °C over 24 hours). »

2.2 Presentation–Standard nomenclature

Any physicist or engineer with moderate training in laboratory instrumentation should be able to look into the flow diagram of a manuscript in AMT or any other equivalent journal and get a basic idea of the method described in it. It is a very important element of a publication of this type, therefore it is my view that the authors should use standard P&ID nomenclature. The current flow diagram is a collection of coloured boxes from which little can be understood. Part of the text uses the photograph supplied (Figure 3) to explain the system something that confuses the reader even more.

2.2.1 Points to consider

- There are standard P&I symbols for valves, pumps, gauges etc that should be used, while colour should be added only if it aids in better explaining the system and not cosmetically. Standard names for the components should be used. For example three pumps with three different descriptions (pressure pump, pump and picarro pump) are found in the block diagram when they (most likely) are of the very same type ie vaccum pump.

The scheme of Figure 2 has been drawn again as:

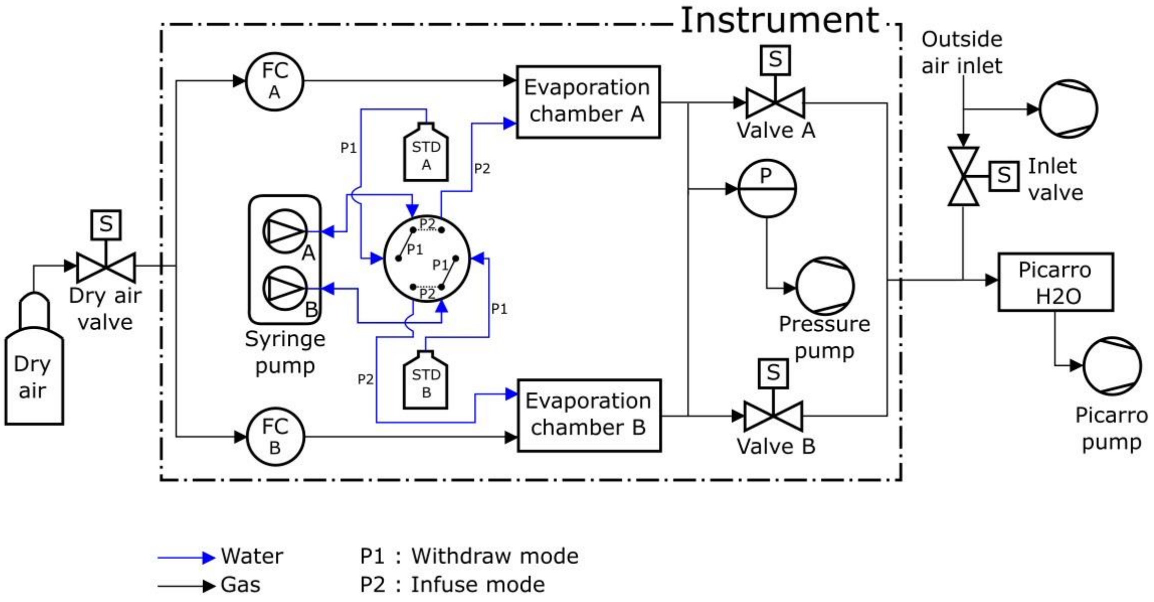


Figure 2: *Humidity generator schematic diagram (see supplementary Table S1 for details on the different elements)*

The same symbol is used for the different pumps but we still give a different name for each pump to help making a distinction between the different pumps in the text.

- One of the main elements of the paper according to the authors is the “double 3-way valve” added to the system. The valve’s type is a 6-port switching valve and a standard symbol for it exists that should be used in this manuscript. It is central to describing the cycles of the system. Feel free to use the supplied (arguably not perfect) Fig. 1 that I am including and describes the valve operation for the syringe filling position.

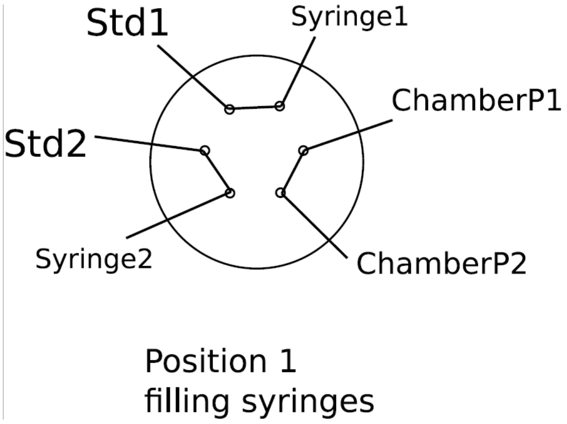


Figure 1

As suggested by the reviewer, we now use the 6-port valve symbol in the figure 2.

- Flows and pressure set points for the regulators also belong to the P&I diagram and since the authors claim that they are important to the operation of the system should be given. A table with the ID numbers of the control units and their set points would be very much appreciated. Currently Table 1 blends names of components from Figure 2 and Figure 3 so the reader has to guess. This is not informative.

In order to better document the system, we propose to add the following tables S1 and 2.

Instruments	Notation on Figure 2	Setting points	Accuracy
Vögtlin GSC-A9TS-DD22	FC	300 and 150 sccm	3.3 sccm
Harvard Apparatus Pump 11 Pico Plus Elite Dual	Syringe pump	0.01 to 0.3 μL/min using 100 μL syringes	0.35% of the set speed
Hamilton syringes	A and B	100 μL	
Swagelok Ultra-Torr SS-4CD-TW-25	Evaporation chamber A and B	Internal volume of 25 cm^3	
Bronkhorst P-702CV-1K1A-AAD-22-V	P	650 to 950 mbar	3 mbar
KNF N86KNDC 24V	Pressure pump		

Table S1: *Description and setting points of the elements composing the LHLG*

Humidity (ppmv)	Infusion rate (μL/min)	Dry Air flow (sccm)
80	0.01	300
160	0.02	300
320	0.04	300
800	0.1	300
1200	0.15	300
1600	0.2	300
2400	0.3	300

Table 2: *Setting points for water infusion rate and dry air flow at a temperature of 20°C.*

2.3 Experimental System–Explanation of operation
Section 2.2 of the manuscript contains information on the principle of operation. One of the puzzling pieces of information in this section is the carrier gas flow. Following the block diagram and the information in Table 1, we see that for the "Drying Mode", the Flow Controllers A and B sustain each a 400 sccm1 flow through the system. The block diagram does not indicate any open splits and for this particular mode the inlet valve is closed and the exhaust vacuum pump is disabled. This is an immense flow for the standard of a picarro spectrometer that normally can sustain its cavity pressure with sample flows in the order of 30-100 sccm1.

Similarly, during the injection of Standard A or Standard B, the flow is in the order of 450 sccm⍰1, a very high flow level. In this mode though the exhaust vacuum pump is enabled, however we see that the lines of Standard A and Standard B are connected. Thus in Standard A mode, valve B is closed and the generated vapour from standard B, hopefully is evacuated via the exhaust vacuum pump. Is the pressure of 905 mbar enough to make sure that all the 150 sccm⍰1 of standard B are off the system and not mixed with the vapour from Standard A? Why was this value chosen and what kind of tests assure that the two lines are never mixed even though physically connected. The safe way to proceed here would be to simply isolate the two channels, remove the pressure regulator and the vacuum pump and simply install two 3-way valves in the position of valves A and B (which I assume are On/Off???) on which one port goes to the instrument inlet and the other is simply open to the atmosphere in an open split configuration.

2.3.1 Points to consider

• A more thorough description of the flow path is needed and the issue of possible cross contmination between Standard A and B should be addressed.

The instrument has been designed with tubing long enough to prevent any mixing and the issue of mixing has been extensively tested during conception by comparing the isotopic composition of the produced water vapor when only one way was in operation (bottle A filled with standard and bottle B empty) and when two ways were in operations (bottles A and B filled with standards of different isotopic composition). The isotopic composition produced on way A was always the same.
To double check this effect when receiving this review, we performed an additional comparison:

- Case A: bottle B was filled with FP5 standard (δ^{18}O = -48.33 ‰) and bottle A filled with EPB standard (δ^{18}O = -6.24 ‰). Measured δ^{18}O on way B was -48.96±1‰ (humidity 768±5 ppmv)
- Case B: bottle A and B were filled with the FP5 standard. Measured δ^{18}O on way B was -49.05±1‰ (humidity 765±4 ppmv).

There is thus no isotopic difference between the two cases both for δ^{18}O and δD which confirms that there is no mixing in the system. These tests and the absence of mixing effect now constitute text S1 in the revised version of the manuscript:

Text S1: No mixing between standards A and B during vaporization

The instrument has been designed with tubing long enough to prevent any mixing. Still, the issue of mixing has been extensively tested during conception by comparing the isotopic composition of the produced water vapor when only one way was in operation (bottle A filled with standard and bottle B empty) and when two ways were in operations (bottles A and B filled with standards of different isotopic composition). The isotopic composition produced on way A was always the same. An additionnal comparison is displayed below.

- Case A: bottle B was filled with FP5 standard (δ^{18}O = -48.33 ‰) and bottle A filled with EPB standard (δ^{18}O = -6.24 ‰). Measured δ^{18}O on way B was -48.96±1‰ (humidity 768±5 ppmv)
- Case B: bottle A and B were filled with the FP5 standard. Measured δ^{18}O on way B was -49.05±1‰ (humidity 765±4 ppmv).

There is no difference between cases A and B which confirms that there is no mixing in the system.

- The flows are very high - How does the picarro cope with this condition?

There is a waste line on this instrument (corresponding to the outlet of the pressure pump) so that the Picarro laser analyzer is only pumping what it needs through valves A and B as it does during the analysis of atmospheric air through inlet valve. As a consequence, there is no problem for the system. This is now mentioned in the revised version of the manuscript.

“When the instrument is connected to the infrared spectrometer, the excess humid air flow is exhausted to the room through the pressure pump and the spectrometer only pumps what is required (Figure 2).”

- There is no information on the volume of the evaporation chambers and no mention regarding the temperature of the system.

Volume of the evaporation chamber is 25 cm^3 (see Table S1). Temperature is around 20°C (see above and in section 2.2).

- There are references to a two standard calibration protocol in this section. It is not exactly clear if this concerns some water concentration correction calibration or a linear slope SMOW-SLAP calibration–see specific comment on the notable absence of any reference to SMOW-SLAP below.

Indeed, this was missing in the previous manuscript. All the lab-standards used in this study are indeed calibrated versus SMOW using a 3-point calibration which will be mentioned in the revised version of the manuscript. Then, we used two ways for our generator (one way for each lab-standard) in order to perform two-point calibrations of the laser analyzer when it is used to measure water vapor isotopic composition at low humidity. We hope that the answers provided for the other comments below now explain better this aspect. The calibration protocol of our data is now explained with the addition of equations 1 and 2 as well as new section 3.5 given later in the answer to reviewer 2.

- I assume that the system is also used in order to perform a SMOW-SLAP calibration. How are evaporation/fractionation effects in the standards’ vials handled? Based on Figure 3 we see no precautions concerning this.

Thank you for this comment. Indeed, this was not mentioned in the manuscript. Fractionation effects have been tested by sampling the water from the flasks every 2 weeks. We observed a maximum evolution of the isotopic composition of the standard in the bottle by 0.05 ‰ for δ^{18}O and 0.5 ‰ in δD within 2 months in case of intense use of the instrument. This is now mentioned in the new version of the manuscript.

“The water in the water reservoirs is sampled every month to check its isotopic composition and renewed when the level of water is below half the maximum level. A maximum evolution of the isotopic composition of the lab-standard filling the water reservoirs has been observed as 0.05‰ and 0.5‰ respectively for δ^{18}O and δD over a 2-month period. “

2.4 Evaluation–Measurement Stability and precision
The manuscript lacks important information on the evaluation of the performance of the system. Section 3.1 assures the reader that no fractionation is observed during the

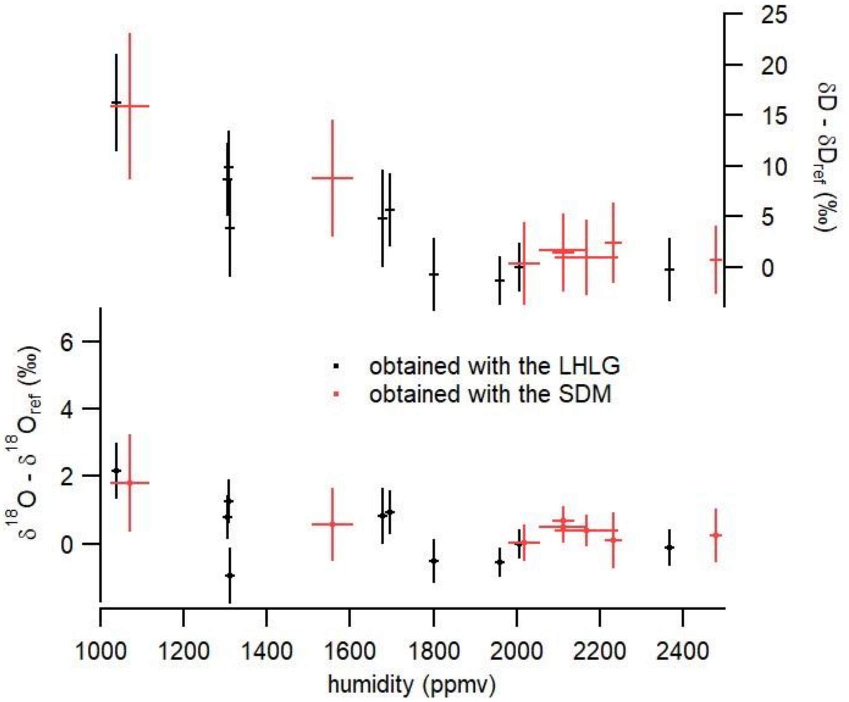
generation of the vapour and its transfer in the flow lines while the agreement with the commercial standard delivery module is within 0.2 and 1‰ for _18O and _D respectively. It is almost impossible to judge these discrepancies based on figure 5 and in fact when looking closely it is rather difficult to see how the differences between the two systems are of this order.

With respect to figure 4 and the evaluation of the precision for [H_2O] and _18O/_D I strongly believe that the manuscript should include a proper Allan variance test (Werle et al , 2011; Steig et al , 2014). It is obvious that the system experiences drifts, whose origing is the vapour generator. Calculating the standard deviation on a 10 min window of a time series that obviously has a changing mean value looks and feels like possible cherry-picking.

2.4.1 Points to consider–Suggestions

- Show better evidence of zero fractionation. If you choose the SDM as a measure of comparison show a detailed comparison and plot the results clearly. The axes of fig 5 span 300 ‰ for _D. It is impossible to say anything.

A focus on the comparison between the SDM and the low-humidity level generator is now displayed on Figure S1:



***<u>Figure S1</u>**: Comparison of the difference in isotopic composition (δD top, $\delta^{18}O$ bottom) between the measured and the reference (measurement performed at 2000 ppmv) values as obtained with the SDM (red) and with the LHLG (black) coupled to the same Picarro L 2130-i with the same lab-standards (FP5) calibrated against VSMOW. The same measured $\delta^{18}O$ and δD values are obtained at 2,000 ppmv through the SDM and the LHLG set-up.*

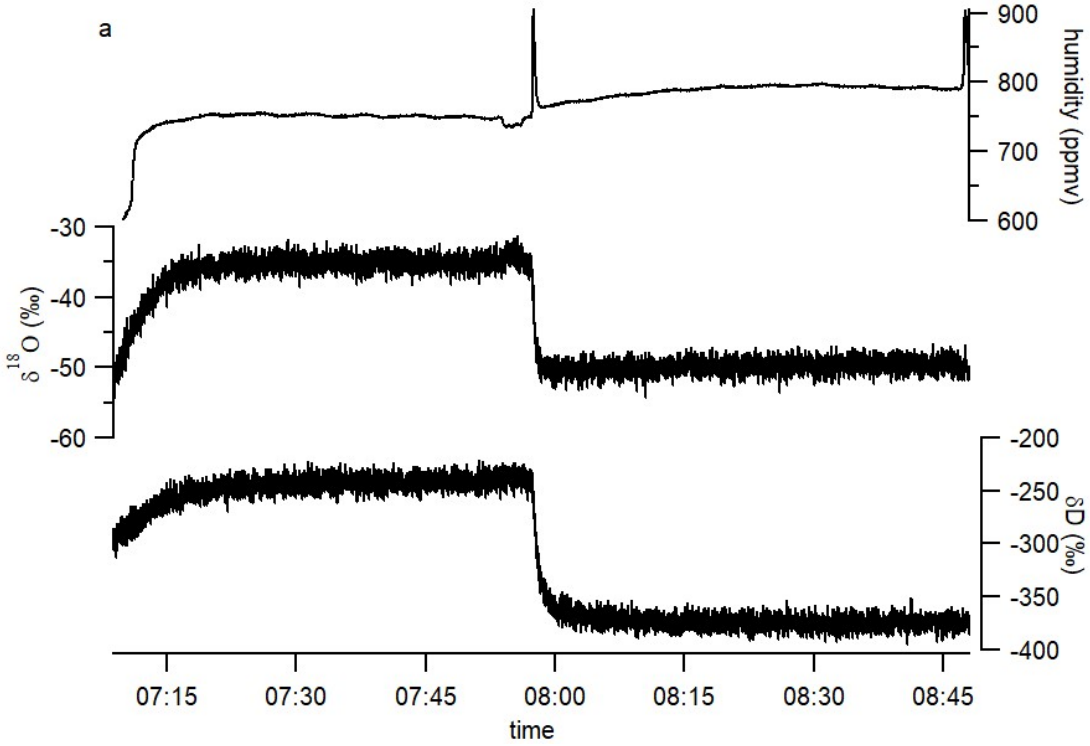
- Calculate the Allan variance of the system for 4-5 [H_2O] levels. Let one water run for several hours and so you get a more honest idea about precision and drifts for all three parameters under consideration. It is a very simple experiment that can provide a valuable insight and by looking in the bibliography it is a rather standard evalluation tool for laser spectroscopy based measurements.

Thank you for the suggestion. Indeed, this is a good idea and we have included the figure 3 (see answer to reviewer 1) in the revised manuscript

2.5 Water concentration correction

The water concentration correction is described in section 3.3 with Figure 5 being the main source of information for this part of the analysis. I would consider this one of the most important sections in the paper and unfortunately it spans only one paragraph. It is repeated in the manuscript that the system presented here is superior to the commercial standard delivery module below the level of 500 ppm. How can we see this when the SDM measurements in this are of [H_2O] are not presented (measurements stop at [H_2O]_2000 ppm)?

It is very difficult to get stable humidity plateau with a SDM below 500 ppmv. This is the reason why we could not display them. However, we include the following figure S2 (in addition to figure S1 shown in the previous answer to comment) in the new manuscript to strengthen the comparison between SDM and LHLG.



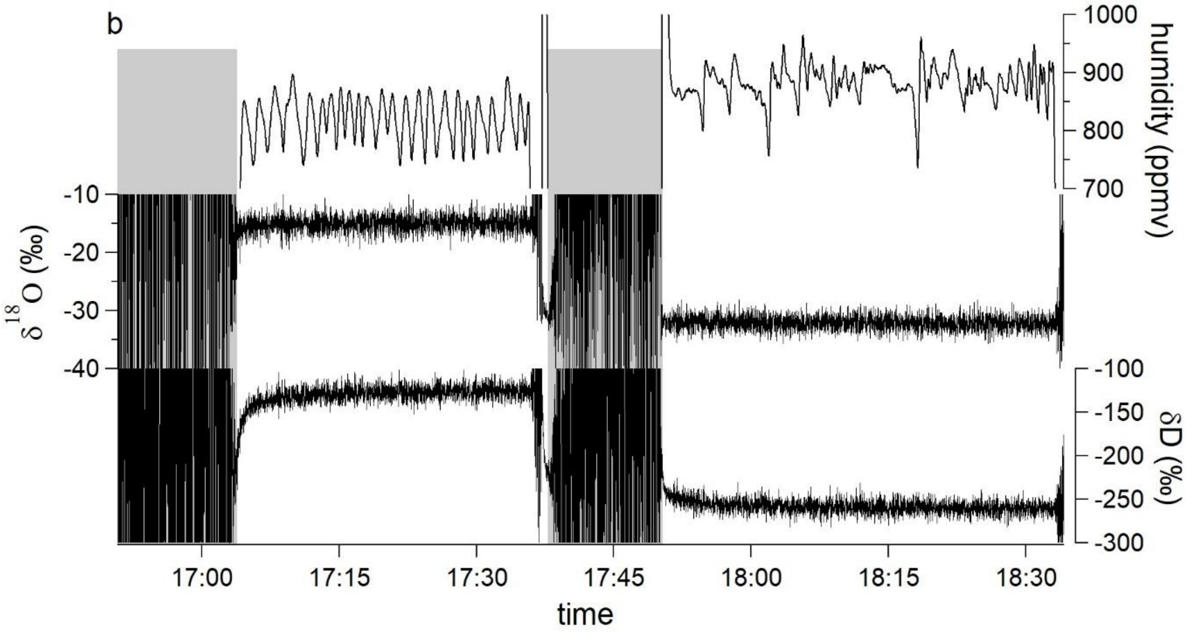


***Figure S2**: Comparison of humidity plateaus (800 ppmv) generated with the LHLG (a) and with the SDM (b). The grey rectangles indicate period with only dry air injected.*

A text has also been added:

"The performance of the present LHLG can be compared to the performance of the SDM (see Supplementary Figures S1 and S2). First (Figure S2), a comparison has been performed at a humidity level of 800 ppmv, for which we have numerous daily calibrations performed with a SDM from a 4.5 years field deployment in Svalbard (Leroy-Dos Santos et al., 2020). The best SDM performance displays a standard deviation 1σ of 31 ppmv, which is significantly worse than the performance of the LHLG (standard deviation 1σ lower than 10 ppmv on average and down to 2 ppmv for 30% of the generated humidity plateaus). Second (Figure S1), while we measure the same influence of humidity on measured $\delta^{18}O$ and δD either with the SDM or with the LHLG, the 1σ values on humidity levels are much larger for the SDM than for the LHLG."

The dataset in Figure 5 presents the difference of the raw isotopic value and a value that the authors call "reference" and "real". These words commonly refer to very specific things in isotope geochemistry and my guess is that the reference value is a SMOW-SLAP calibrated value (?). The term SMOW-SLAP calibration is not to be found at all in the manuscript. It is to some extent problematic that a water concentration correction measurement compares a raw value at a number of [H_2O] levels with a value post SMOW-SLAP calibration. Please see my comments on SMOW-SLAP in the next section.

As mentioned above, the SMOW-SLAP calibration is now addressed in the revised manuscript and all our lab-standards are calibrated vs V-SMOW (using also V-SLAP as a reference).

As an example:
"First, the isotopic composition of three different lab-standards calibrated against VSMOW at LSCE (H_2O-CO_2 equilibration followed by IRMS for $\delta^{18}O$; Cavity RingDown Spectroscopy for δD; calibrated every 3 years using VSMOW and VSLAP provided by IAEA) have been compared, after their

generation by the present LHLG and by the commercial SDM, both at a humidity of 2,000 ppmv over 50-min time spans."

The type of correction is not described in the manuscript. This is what this work is all about. There can be various approaches on how it can be done (see for example (Schmidt et al , 2010; Gkinis et al , 2010; Aemisegger et al , 2012)) but the authors need to be open and specific about what they did exactly. Moreover (Schmidt et al , 2010) suggest that the response to changing [H_2O] levels varies with the isotopic composition. I do not know how important this effect is for this present work, but I believe that the authors should perform two separate experiments with two different waters (technically the system offers this possibility). It would also strengthen their claims for a vary stable system if they show the full experiment with its raw data as a function of time and not only the averages.

There is also a claim in this section that the stability of the system allows via the daily measurement of standards waters to quantify instrumental drifts. There is no strict evaluation of the stability of the system so far and I think that a proper Allan variance test at various levels of [H_2O] is necessary. But even in the absence of an Allan variance test the data given in Table 2 do not necessarily indicate instrumental drifts. All the isotope data in the table are within 1 ⍰ _ of the noise level for a 10-min period mentioned in the caption and there does not seem to be a one way trend. So how is it possible to say anything about instrumental drifts? Also the noise levels mentioned in the caption of Table 2 are very different for roughly the same [H_2O] when one looks at the data of Figure 4 (4.5 versus 14).

This is a good remark indeed and we are sorry for the confusion. Actually, while the noise on the humidity level is mainly due to the low-humidity level generator, the noise on the isotopic ratios depends on the laser analyzer which explains some inconsistency in the results which were displayed (we presented results obtained with different instruments). The inconsistencies will be fixed in the revised version. Indeed, Table 2 shows that there is no measurable drift and this is now explained in the revised version.

"The stability of the LHLG allows a robust quantification of the L2130-i analyzer drift thanks to a daily measurement of the same water isotopic standard reference (see Table S3 showing actually no measurable drift over a 3-week period)."

2.5.1 Points to consider-Suggestions
- A clear description of the correction is missing. Some math is needed here.

We provide the following explanation in the revised manuscript.

Our data show a result already observed in Weng et al. (2020): while the dependency of δ^{18}O and δD to humidity is similar for low δ^{18}O and δD lab-standards (NEEM and FP5), we observe a different behavior for the δD vs humidity relationship for the high δ^{18}O and δD lab-standard EPB. This result strengthens the recommendation of Weng et al. (2020) to use two water standards in the range of the measured water vapor isotopic composition to best calibrate our final data. In our case, our applications were in Antarctica, so that we used our two lowest lab-standards (NEEM and FP5). For the two standards and for this particular Picarro L2130-i (results are expected to depend on the instrument), the same dependency of isotopic composition vs humidity is observed. We express this dependency as the relationship between the difference in δD or δ^{18}O between the measured value at the given humidity and the value of the same standard measured at a humidity of 2,000 ppmv. The experimental data for NEEM and FP5 from Figure 5 are fitted through polynomial functions with respect to humidity h (in ppmv):

$\delta^{18}O - \delta^{18}O_{ref} = 3.97\times10^{-18}\times h^6 - 3.59\times10^{-14}\times h^5 + 1.28\times10^{-10}\times h^4 - 2.31\times10^{-7}\times h^3 + 2.19\times10^{-4}\times h^2 - 1.06\times10^{-1}\times h + 23.7$ (eq.1)
$\delta D - \delta D_{ref} = 6.86\times10^{-17}\times h^6 - 6.00\times10^{-13}\times h^5 + 2.08\times10^{-9}\times h^4 - 3.61.10^{-6}\times h^3 + 3.31.10^{-3}\times h^2 - 1.54\times h + 313$ (eq. 2)

After this correction, the measured values corrected from humidity dependence are corrected using the comparison of the measured values of the 2 standards at 2,000 ppmv to their VSMOW calibrated values as explained in section 3.5 below."

- An experiment at two different isotopic levels showing the raw data versus time will show if there is an isotope effect in the water concentration dependence.

We think that new figures 3 and 5 answer exactly this comment.

- A proper quantification of instrumental drifts (and this concerns the system as a whole and not only the spectrometer) can be done with a proper Allan variance test.

This is now shown on figure 3 which is implemented in the revised version of the manuscript.

2.6 Lack of reference to SMOW-SLAP calibration

The main goal of building a water vapour generator as a peripheral for isotope measurements of water vapour is to be able to calibrate the dataset on the SMOW-SLAP scale. This is the only way to communicate and compare the measurements with other existing data sets and produce some science out of them. It is also even more important if the deuterium excess parameter will be studied as it is very sensitive to this calibration procedure. Therefore it appears very awkward that a manuscript dealing with this topic does not include a single note, comment or reference to this very important step Of the measurement process. The dataset presented later in section 4 of the manuscript are impossible to evaluate if they are not calibrated in the SMOW-SLAP scale.

One more purpose of performing such calibrations, is that they can reveal possible accuracy issues in the instrumentation system. Given two standard waters one should be able to produce a calibration line and thereafter measure a third water of known isotopic composition treating it as an unknown. If the resulting value lies beyond the 3 ⍰ _ range then there is likely something wrong with the system. That could be any part from the water standard storage to the water vapour generation system or the spectrometer itself. Currently there is no way to say anything about the accuracy of the system. With this in mind, section 4 of the manuscript is of very little use as the dataset is reported on some local instrument scale.

A SMOW-SLAP calibration experiment at various [H_2O] levels using the SDM and the current system would provide a proper comparison between the two systems and therefore it would be a very important addition to the manuscript.

2.6.1 Points to consider

- A proper treatment of the SMOW-SLAP calibration step is notably missing.

All our lab-standards are regularly calibrated vs SMOW using IAEA VSMOW and VSLAP standards (every 3 years in minimum). In order to calibrate our lab-standards, we use determination from

Picarro L 2130-i for δD and for $\delta^{18}O$. In addition, we use a $\delta^{18}O$ calibration using an IRMS through water – CO_2 equilibration.
New text:

"First, the isotopic composition of three different lab-standards calibrated against VSMOW at LSCE (H_2O-CO_2 equilibration followed by IRMS for $\delta^{18}O$; Cavity RingDown Spectroscopy for δD) have been compared, after their generation by the present LHLG and by the commercial SDM, both at a humidity of 2,000 ppmv over 50-min time spans."

- Performing 2-standard calibrations and measuring a third water standard treated as an unknown will be a valuable -almost essential- addition to the manuscript, offering important information on the accuracy of the system.

This is an excellent suggestion and this test has been performed and a new section 3.5 has been added:

3.5- Accuracy of the system

The accuracy of the system has been addressed performing a 2-standard calibration and measuring a third standard treated as an unknown. We used two lab-standards calibrated vs VSMOW with large $\delta^{18}O$ and δD differences (EPB and FP5) and used the lab-standard NEEM, also independently calibrated against VSMOW. The 3 lab-standards have been vaporized at 800 ppmv and measured by the same L2130-i analyzer.

Standard	VSMOW calibrated value	Measured value at 800 ppmv	Measured value corrected from humidity dependence (Equation 1)
EPB	-6.24 ‰	-8.27 ‰	-10.78 ‰
NEEM	-33.5 ‰	-34.48 ‰	-36.99 ‰
FP5	-48.33 ‰	-49.02 ‰	-51.53 ‰

Table 3: *Comparison of measured vs VSMOW calibrated $\delta^{18}O$ values for 3 standards measured with a Picarro analyzer after generation of water vapor using the LHLG.*

We used the measured and true values of EPB and FP5 to estimate the $\delta^{18}O$ value of the NEEM standard from its measured value (Table 3). Using the linear relationship obtained from VSMOW calibrated EPB and FP5 $\delta^{18}O$ vs measured EPB and FP5 $\delta^{18}O$ values following the recommendations of the National Institute of Standards and Technology (NIST, reference material 8535a) leads to an estimated NEEM $\delta^{18}O$ of -33.31 ‰ to be compared to the independently VSMOW calibrated value of -33.5 ‰. Given the uncertainty of about 0.8-1 ‰ when measuring $\delta^{18}O$ around 800 ppmv, we can conclude that the system is accurate.

- Since a lot has been written about the performance of the commercial SDM it will be proper to perform 2-standard calibrations for various [H_2O] levels and compare the results.

See answer above as well as the Figures S1 and S2 added to answer this comment.

References:

Kerstel, E.: Modeling the dynamic behavior of a droplet evaporation device for the delivery of isotopically calibrated low-humidity water vapor, Atmos. Meas. Tech. Discuss. [preprint], https://doi.org/10.5194/amt-2020-428, in review, 2020.

Landsberg, J.: Développement d’un spectromètre laser OF-CEAS pour les mesures des isotopes de la vapeur d’eau aux concentrations de l’eau basses. [online] Available from: http://www.theses.fr/2014GRENY052/document, 2014.

Merlivat, L., Molecular diffusivities of H216O, HD16O and H218O in gases, J. Phys. Chem., 69(6), 2864–2871, 1978.