Simple water vapor sampling for stable isotope analysis using affordable valves and bags

Adrian Dahlmann[1], John D. Marshall[2], David Dubbert[1], Mathias Hoffmann[1] and Maren Dubbert[1]

[1]Isotope Biogeochemistry and Gas Fluxes, Leibniz Centre for Agricultural Landscape Research (ZALF), Müncheberg, 15374, Germany
[2]Department of Earth Sciences, University of Gothenburg, Gothenburg, 405 30, Sweden

Correspondence to: Adrian Dahlmann (adrian.dahlmann@zalf.de)

Abstract. Water stable isotopes are commonly used in hydrological and ecological research. Until now, most measurements have been made by taking a sample from the field and extracting its water in the laboratory. More recently, samples have been collected with semi-permeable membranes and measured in the field. These new methods, however, present challenges in achieving high-resolution measurements across multiple sites since they require significant effort and resources. Gas bag sampling offers the advantage of non-destructive, cost-efficient, easy-to-perform measurements without the need to bring a Cavity Ring-Down Spectroscopy (CRDS) analyzer into the field. We used gas-permeable membranes (GPM) to extract samples of water vapor from the soil, which were then stored in multi-layer foil bags until analysis. The bags were modified with home-made connections to reduce leakage and simplify gas transfers. The bags were tested using laboratory standards to determine their maximum storage time, potential memory effects, and reusability. The storage experiment demonstrated the ability to store water vapor samples for up to seven days while maintaining acceptable results for δ^2H and δ^{18}O. The memory experiment showed that the influence of previous samples increased with duration of storage. The reuse experiment showed that samples can be filled repeatedly, provided they are filled and evacuated ten times with dry air between measured samples. To demonstrate bag applicability in the field, we compared accuracy and precision of stored samples to measurements made directly in the field. Trueness was defined as mean difference between the measured and known water vapor placed into the bags and precision by the standard deviation of repeated measurements. With new gas bags, we found accuracy of 0.2 ± 0.9 ‰, respectively, for δ^{18}O and 0.7 ± 2.3 ‰ for δ^2H. When the bags were reused, they yielded accuracy of 0.1 ± 0.8 ‰ for δ^{18}O and 1.4 ± 3.3 ‰ for δ^2H. This is at least as good as published alternative methods. The proposed system is simple, cost-effective, and versatile for both lab and field applications.

1. Introduction

Water stable isotope measurements are used in a variety of scientific fields, particularly in hydrology, ecohydrology, and meteorology, which focus on aspects of the water cycle. The primary isotopes involved are ^{18}O and ^{2}H (e.g., Gat 1996; Mook 2000), described as δ^{18}O and δ^{2}H relative to the most abundant isotopes, ^{16}O and ^{1}H (Sodemann, 2006). They serve to investigate processes such as infiltration and groundwater recharge (e.g. Séraphin et al., 2016), evaporation (e.g. Rothfuss et al., 2010), or the plasticity of root water uptake under stress (e.g. Kühnhammer et al., 2021; Kühnhammer et al., 2023).

Traditionally, the isotopic composition of soil and plant water has been measured through destructive sampling of soil cores or sampled plant material, followed by water extraction e.g. via cryogenic extraction (see method summary Orlowski et al., 2016a) and measured with isotope ratio mass spectrometry (IRMS) analyzers (West et al., 2006; Sprenger et al., 2015). The development of smaller and less expensive cavity ring-down spectroscopy (CRDS) analyzers has led to an increase in potential applications, including, e.g., in situ measurements using gas permeable membranes (Rothfuss et al., 2013; Volkmann and Weiler, 2014; Volkmann et al., 2016; Kübert et al., 2021; Marshall et al.; 2020). Direct measurements are a viable alternative to classic destructive techniques, especially in small plots, as among other benefits (e.g. high frequency measurements) they avoid repeated destructive sampling. However, direct, continuous in situ field setups are very cost-intensive and technically challenging, requiring a laser spectrometer (e.g. a CRDS) and permanent power supply in the field as well as a strong expertise to maintain. To allow an expansion to a wider set of potential study areas and increase the number of absolute study areas maintainable, scientists are recently trying to develop new simplified sampling systems. This includes capturing soil moisture as water vapor for subsequent laboratory analysis (e.g. Havranek et al., 2020; Magh et al., 2022; Herbstritt et al. 2023). To do so, primarily glass bottles or gas sampling bags with various fittings are used, which cost from ~1-200 euros per container. The advantages of these methods include the ability to quickly measure stored samples in a temperature-stable laboratory environment. In addition, multiple sample containers can be filled at once in the field, which allows for the simultaneous measurement of multiple probes, and sampling can generally be performed at a much faster rate. These simplified and more affordable systems could therefore increase the number of studies on water stable isotopes and provide new insights in research by increasing the number of possible experimental sites and samples.

In this study, we investigated the use of multi-foil bags with septum valves. Our investigation focused on exploring the potential of these commercially available but affordable bags for a wider range of applications (~ 20€ per bag plus ~ 15€ for the connection). To ensure easy and reliable bag filling and measurement, we built an additional connection and a portable dry air supply box system for easy field measurement. We tested the prepared bags in several experiments in the laboratory using defined standards and, in the field, using comparison to in situ measurements with a CRDS. These results allowed us to find a simple approach to using septum-based gas bags for field measurements of water stable isotopes. The focus was to investigate storage capability as well as possible isotopic fractionation effects due to exchange with the inner surface of the bags. Five different experiments were performed: i) a storage experiment up to 7 days, ii) a memory experiment without sample storage and two quite different standards, iii) a memory experiment with one day of storage of the initial standard followed by sample replacement exploring duration effects on memory setting and, iv) a field filling and bag reuse experiment to compare the bag measurements with in situ CRDS measurements. These were followed by v) a gas bag measurement sequence over a full cultivation period.

2. Material and methods

2.1 Study area and basics of water stable isotope measurements

The laboratory experiments were carried out at the Leibniz Centre for Agricultural Landscape Research (ZALF). The field experiments took place at the AgroFlux experimental platform of ZALF, located in the northeast of Germany, near Dedelow in the Uckermark region (N 53°22′45", E 13°47′11"; ~50-60 m a.s.l.).

During the experiments, the δ^2H and δ^{18}O values were recorded using a CRDS analyzer (L2130-i, Picarro Inc., Santa Clara, CA, USA). The hydrogen and oxygen stable isotopes in the sampled water vapor (δ^2H and δ^{18}O) are given in per mil (‰), relative to the Vienna Standard Mean Ocean Water (VSMOW) using δ-notation (Eq. 1; Craig, 1961).

$$\delta = \left(\frac{R_{sample}}{R_{VSMOW}} - 1\right) \times 1000 \qquad \text{Eq. 1}$$

During all experiments, water stable isotope signatures (δ^2H and δ^{18}O in ‰) were measured with the method of Rothfuss et al. (2013), using gas permeable membranes (GPM, Accurel GP V8/2HF, 3M, Germany; 0.155 cm wall thickness, 0.55 cm i.d., 0.86 cm o.d.; e.g. as used in Kübert et al., 2020 or Kühnhammer et al., 2021). In the laboratory experiments, we attached two membranes to the cap of a 100 ml glass bottle with two stainless steel fittings (CUA-2, Hy-Lok D Vertriebs GmbH, Germany) to directly measure standard water vapor and to fill the bags.

The glass bottle was filled with approx. ~ 60 - 80 ml of standard water. The first membrane was submerged in the standard water, where it bubbles the dry air through, resulting in equilibration of water vapor in the headspace with the standard water. The second membrane, in the headspace, collects saturated sample air and supplies it to the analyzer. Both membranes were sealed with adhesive. The second membrane (< 5 cm) served as a safety mechanism to prevent liquid water from entering the tubing.

A gas cylinder was used to induce dry gas at a low flow rate of 50 - 80 ml per minute (257-6409, RS Components GmbH, Germany). We ensured that the isotopic signature of the vapor would be at equilibrium with liquid water at this flow rate. We tested flows from the minimum required for Picarro operation (approximately 35 ml/min) to 300 ml/min and found accurate results to 100 ml/min. At the lower flow rates, the water vapor passing through the membrane reached isotopic thermodynamic equilibrium (Majoube, 1971; Horita and Wesolowski, 1994). In the field experiments, we used approx. 12 cm membranes (comparable to soil GPM in e.g. Kühnhammer et al., 2021) attached to PTFE tubing to sample the four different soil depths (see section 2.7). The in situ method was likewise based on the measurement of water vapor with the assumption that the vapor was in isotopic equilibrium with the liquid water surrounding the sample probe (Rothfuss et al., 2013). Finally, the isotopic fractionation was calculated as a function of the temperature (T) at the phase transition using equations based on Majoube (1971).

The water vapor from the standards and soil was then either transferred immediately to the CRDS analyzer and measured directly or it was stored in the gas bags and measured later. In laboratory experiments I, II and III, the temperatures were around 20°C during filling and around 24°C during storage and bag measurement to avoid condensation. In field experiments IV and V, great care was taken to measure the bags at elevated temperatures relative to the source temperatures.

2.2 Storage and sampling design

2.2.1 Gas bag design

The sampling and measurement concept was intended to be as simple as possible, while still providing high accuracy and precision. Water vapor samples were stored in 1-L multi-layer foil bags with a septum-based valve (1l Multi-Layer Foil Bags with stainless steel fitting, Sense Trading B.V., Netherlands; see Table S1 for more details; Sense Trading B.V., 2024). The stainless steel 2-in-1 fitting combined the valve and septum, with the septum acting as a seal, allowing air to flow around it when the valve was open, and sealing when the valve was closed.

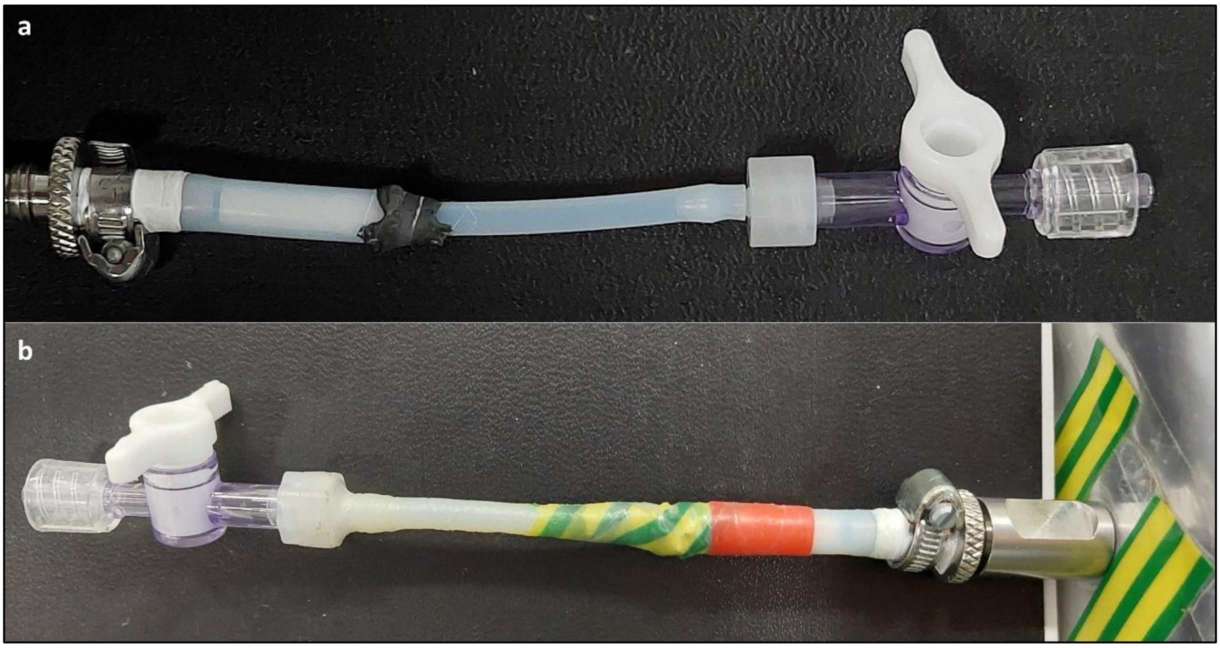

Figure 1: Self-constructed luer-lock connector with the splice exposed (a) and stabilized with tape attached to the bag on the right (b).

As recommended by the manufacturer, care was taken when filling the bags to ensure that the maximum volume did not exceed 90% of capacity, which could cause material damage. The connection (Fig. 1) was built to easily attach the bags with the sample setup. It consisted of two short PTFE tubes (PTFE-tubing, Wolf-Technik eK, Germany) and an additional luer-lock stopcock (1-way Masterflex™ Stopcocks with Luer Connection, Avantor, USA). A hose clamp (TORRO SGL 5mm, NORMA Group Holding GmbH, Germany) was used to directly connect a quarter-inch tube to the valve and the other 4 mm tube was glued into the quarter-inch tube using 2-component-adhesive (DP8005, 3M Deutschland GmbH, Germany). Since the adhesive contact with the PTFE tube could break under tension and cause leakage, we wrapped electrical insulation tape around the splice to reinforce the connector. This tape was not essential for sealing. Then, a luer-lock connection (LF-1.5NK-QC, GMPTEC GmbH, Germany) was used to connect the luer-lock stopcock.

2.2.2 Sampling design

For the 1) direct standard measurements, the sample thus generated was passed directly to the laser spectrometer to determine its isotopic signature. Since the laser spectrometer only has a flow rate of approx. 35 to 40 ml per minute, an open split was added to ensure a constant flow and to avoid pressure differences. The open split was measured continuously to ensure that no ambient air could flow back. A 5-minute average was taken at the end of a minimum 10-minute measurement for direct standard measurements.

For the 2) field measurements, the membranes were installed at the four different depths of 5 cm, 15 cm, 45 cm and 150 cm, and water vapor was transported out of the soil using 4 mm

PTFE tubing. The open ends were fitted with Luer connectors for later connection of gas sample bags and the dry air supply. To protect these open ends from environmental influences, waterproof outdoor boxes were installed 20 to 30 cm above the ground (outdoor.case type 500, B&W International GmbH, Germany). Cable glands were used to keep the boxes watertight (PG screw set, reichelt elektronik GmbH, Germany).

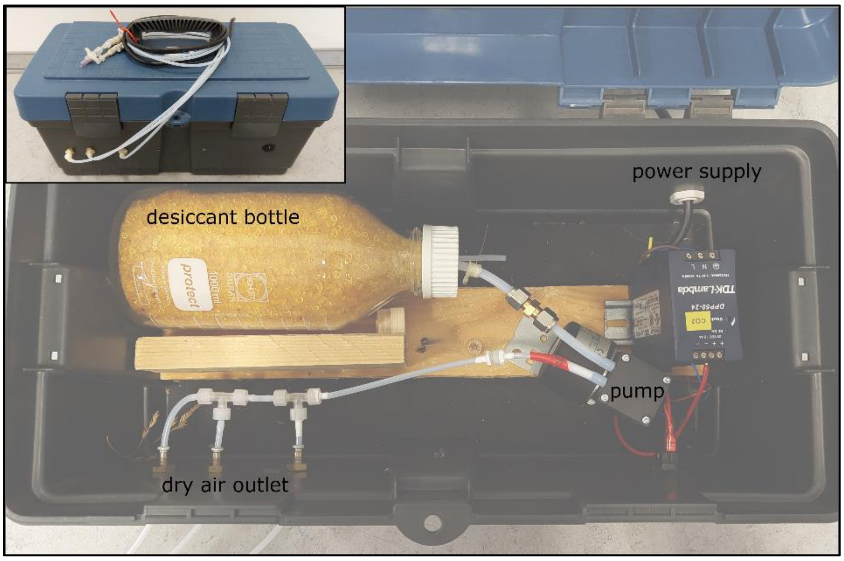

Figure 2: Self-constructed box for field dry air supply (top left) including a bottle with desiccant, power supply and a pump for up to three dry air outlet lines.

A separate box was built to supply pressurized dry air to the measuring system during the field experiments (Fig. 2). This contained a pump (NMP850KPDC-B, KNF DAC GmbH, Germany) including a power supply (DPP50-24, TDK-Lambda Germany, Germany), which could transport the dry air in three tubes simultaneously through up to three sample lines. The air is ambient air which is dried by a desiccant (Silica Gel Orange, Carl Roth GmbH + Co. KG, Germany) contained in a 1-liter bottle (Screw top bottle DURAN®, DWK Life Science, USA). To regulate the flow of individual sample lines, fixed valves were used (AS1002F-04, SMC Deutschland GmbH, Germany). The dry air supply box was tested prior to our experiments by measuring the outlet concentration of the dry box over the course of one day. However, the use of such a system should always be tested for the specific application, as a very high flow rate combined with very humid air could greatly affect the duration of possible use. During the experiments, we periodically tested the water concentration before and after the field campaigns and could not detect any increase after one day in the field. The water concentration of the dry air produced was approximately 200 ppm.

2.3 Laboratory standards

The water stable isotope measurements were calibrated against six water vapor standards (see Table 1) that were manually measured during the experiments. Temperature (T) was recorded

Table 1: Liquid water standards used during the experiments.

Standard	$\delta^{18}O_{liquid}$ [‰]	$\delta^{2}H_{liquid}$ [‰]
L22	- 19.9	- 148.1
M22	- 9	- 63.3
H22	2	12.9
L23	- 16	- 108.2
M23	- 9.2	- 63.9
H23	- 1.3	- 32

continuously every 30 seconds with a thermometer (EBI 20-TH1, Xylem Analytics Germany Sales GmbH & Co. KG, Germany) placed directly next to the standard container. This allowed us to measure the standards in the vapor phase and infer the values in a liquid phase at equilibrium (Sec. 2.5). Of the six standards with different δ values, approximately 60 ml were filled into the prepared 100 ml standard bottles as described in section 2.1 and measured directly on the CRDS.

2.4 Experimental design

2.4.1 Experiment I: Storage duration

In our storage experiment, we tested our gas sample bags for water vapor storage using water sources of known isotopic composition. New bags, including the self-made connections were prepared to eliminate any production artifacts. Each bag was cycled with dry air, filled, and emptied five times in a row. Following this preparation, five bags per storage period were filled with two standards, L22 and M22 (15 min. at 50 ml/min).

Upon filling, the gas bags were promptly measured to ensure that no isotopic fractionation occurred during the filling process. Subsequently, the gas bags were stored in the laboratory for three storage durations - 1 day, 3 days, and 7 days. After the designated storage periods, the samples were measured for 4 to 5 minutes, and a stable 2-minute average was recorded.

2.4.2 Experiment II: Memory

We conducted two memory tests, maintaining a consistent methodology similar to that employed during the storage experiment, both utilizing newly prepared bags. In the first test, we followed a structured sequence: we filled gas bags with this standard, emptied them, and switched to another standard and refilled the bags. We repeated the process (fill, measure, empty) with the opposite standard until our measurements fell within the required accurate range (defined in 2.5). In the first experiment, L23 was used as the initial standard and H23 as the opposite standard, in the second experiment, the standards were used in reverse order. We used five gas bags per standard during the experiments.

2.4.3 Experiment III: Memory test with storage

This laboratory experiment was conceived after we observed the effect of a short delay on memory in Experiment II. We followed a similar procedure except that the initial standard was allowed to stand in the bags for one day prior to replacement with the second standard. We then proceeded with the second standard following the repeated steps (fill, measure, empty) until our measurements fell within the accurate range again. Between the second and third measurement cycle, the experiment was interrupted due to the long duration (1h) of each measurement cycle and continued the next day (after 15.5 hours). The bags were left empty during this second night to avoid any effects. Due to the length of each measurement cycle, we used 3 repetitions during the experiment.

2.4.4 Experiment IV: Field filling and bag reuse

To validate results gained during the laboratory experiments under field conditions, we compared measurements using the gas bags with direct in situ CRDS measurements. To do so, we conducted two measuring campaigns, the first using new bags and the second using reused bags. During the first one, we focused on the applicability of bag filling in the field by comparing direct measurements of the soil water isotopes with the CRDS in the field measurement of bagged samples. In the second campaign, we again compared direct field measurements to bagged measurements, but this time using re-used bags measured in the laboratory within 24 hours. To exclude any memory effects, as we saw in experiment III, the reused bags were rinsed 10 times with dry air (approx. 10 x 10 min). Identical sample bags were utilized for the identical sample probe to minimize changes in isotopic composition and reduce the impact of memory effects. During each of the two measurement campaigns, a total of 48 samples were collected at four different depths: 5cm (n = 14), 15cm (n = 13), 45cm (n = 7), and 150cm (n = 14). Due to low soil permeability, the depth of 45cm could only be sampled during one measurement campaign, resulting in only 7 samples. Dry carrier gas was passed through the home-built membrane soil probes at a flow rate of approximately 50 ml per minute. First, we connected the CRDS to the outlet valve to determine the time required to reach a steady-state value (compared to e.g. Kühnhammer et al., 2021). Subsequently, a 2-minute average was recorded at the end of a 15-minute measurement for comparison with the subsequent bag measurement. Second, we connected the bags and filled them for 15 minutes (approx. 750 mL). The temperature at the sampled soil depth (TEROS 21, Meter Group, USA) was logged using a datalogger (CR1000, Campbell Scientific Ltd., Germany) at 20-minute averages and used to correct for equilibrium fractionation. Furthermore, it was used to determine the saturated water concentration to control the concurrent measured concentration in the probe.

2.4.5 Experiment V: Observation over a full cultivation period

The field applicability test was followed by gas bag sampling and subsequent water stable isotope analyses in the laboratory for the same soil depths during a full winter wheat cropping period (variety: "Ponticus"; sowing: September 26, 2022; harvest: July 18, 2023). We measured once a month during the winter and once a week starting in the spring resulting in 18 measurement campaigns using only our gas bags. For calibration, three laboratory standards were bagged and treated in the same manner as the samples. As was the case with experiment IV, identical sample bags were used for the identical sample probes throughout all campaigns. Sample bags were replaced with new ones if they were damaged. To provide context for the soil isotopic data, additional precipitation samples were collected at the site over a two-year period.

2.5 Calculation of isotope ratios, evaluation of uncertainty and data correction

The isotope signatures of the collected water vapor water sample were converted to liquid water isotope signatures using Majoube's method (Majoube, 1971; Kübert et al., 2020). This conversion was based on equilibrium fractionation at the source temperature (Eq. 2 and 3).

$$\delta_{liquid} = \left(\delta_{vapor} + 1\right) \times \alpha^{+} - 1 \qquad \text{Eq. 2}$$

$$\ln \alpha^{+} = a\,\frac{10^6}{T^2} + b\,\frac{10^3}{T} + c \qquad \text{Eq. 3}$$

The equilibrium fractionation factor $\alpha+$ was determined based on Majoube’s (1971) experimental results, using the coefficients a, b and c (a = 1.137, b = -0.4156 and c = -2.0667 for ^{18}O and a = 24.844, b = -76.248 and c = 52.612 for ^{2}H).

To assess the accuracy of our laboratory measurements, we calculated z-scores for each sample and water stable isotope (δ^{18}O and δ^{2}H). Z-scores indicate the normalized deviation of the measured water isotopic ratios from the known isotopic signature of the added water vapor, and can be calculated following the method (Eq. 4) described by Wassenaar et al. (2012):

$$z - score = \frac{S - B}{\mu} \qquad \text{Eq. 4}$$

Where S is the isotope signature (δ^{18}O or δ^{2}H) measured with our gas bag, B is the benchmark isotope signature and μ is the target standard deviation. To assess the performance of each extraction method, we set a target standard deviation (SD) of 2‰ for δ^{2}H and 0.4‰ for δ^{18}O for measuring water vapor samples. The target SD was selected based on CRDS measurements using the bag method and considering standard deviations from previous studies, such as those by Wassenaar et al. (2012) or Orlowski et al. (2016a). A z-score < 2 represents an accurate sample range, a z-score between 2 and 5 describes the questionable range, and a z-score > 5 representing an unacceptable range (Wassenaar et al., 2012; Orlowski et al., 2016a).

3. Results

The experimental results will be described using the following figure design: the defined standard deviation will be shown as a dashed blue box in plots of the true water vapor isotope values, which will be predominantly shown on the left side. The accurate z-scores are shown as a dashed black box and the questionable z-scores are shown as a black box, predominantly on the right side. Both standard deviation and z-scores were defined in section 2.5.

3.1 Experiment I: Storage duration

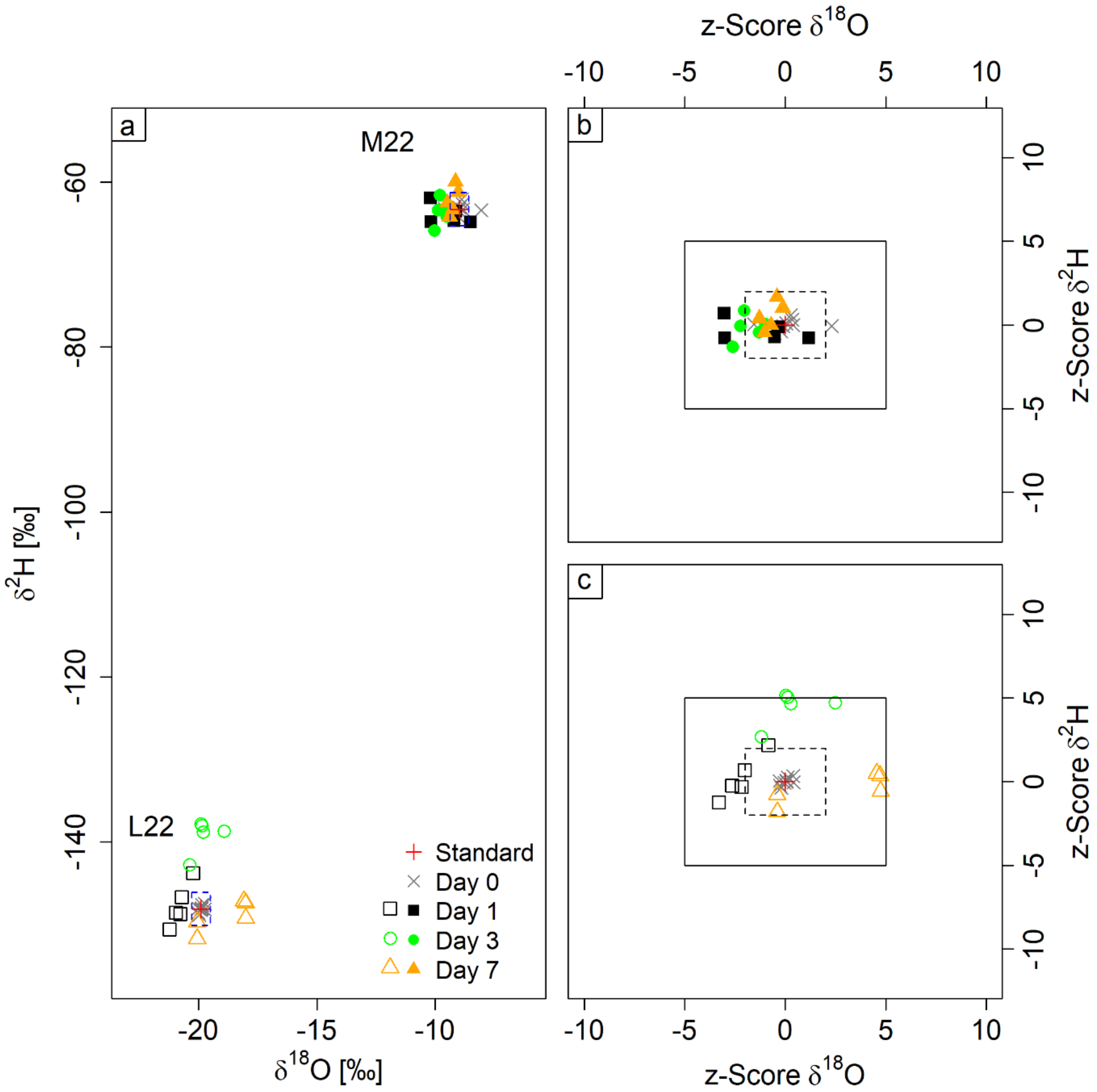

Figure 3: Dual isotope plots showing variation over several days of water-vapor storage in gas bags. The panel on the left shows results from both experiments (a) and those on the right show z-score plots for standard M22 (filled symbols, b) and L22 (open symbols, c). The black boxes describe the questionable range while the boxes delineated with a dashed line describes the accurate range (b, c). The blue dashed line (a) describes the defined standard deviation for measurements.

Used laboratory standards, L22 and M22, spanned an isotopic gradient of – 9.0 to - 19.9 ‰ in δ^{18}O and - 63.3 to - 148.1 ‰ in δ^{2}H (Fig. 3a; filled symbols: M22, empty symbols: L22). On average, measured values differed from known values by -0.7 ± 0.6 ‰ δ^{18}O and -0.1 ± 2 ‰ δ^{2}H after 1 day, -0.3 ± 0.6 ‰ δ^{18}O and +4.3 ± 5.2 ‰ δ^{2}H after 3 days and, +0.4 ± 1 ‰ δ^{18}O and +0.1 ± 2 ‰ δ^{2}H after 7 days of storage (Table S2/S3). Except for one sample during the M22 experiment, deviations from the true standard values in these measurements were all in the range of ± 0.4 for δ^{18}O and 2 ‰ for δ^{2}H and thus no bias was associated with filling of the bags.

The experiment using standard M22 resulted in an overall high accuracy for all measurements of the three storage durations being – 0.5 ± 0.5 ‰ for δ^{18}O and 0 ± 1.6 ‰ for δ^{2}H. In addition, no trend in isotopic signature could be observed over storage duration for either δ^{18}O or δ^{2}H. Consequently, z-scores were either within the accurate range or close to it, again with no trend of decreasing accuracy over storage time.

The second storage test using L22, showed a lower accuracy due to lower precision for δ^{2}H, being – 0.1 ± 1.1 ‰ for δ^{18}O and 2.8 ± 4.9 ‰. However, no time trend was observed. The decreased accuracy was mostly caused by the samples after three days, as all gas bags showed a significant enrichment (8.9 ± 2 ‰ on average). The higher inaccuracy after three days of storage must be due to an error during the measurement, as accuracy improved again after 7 days. The overall higher scatter (particularly for δ^{18}O), which has a different isotopic signature than the ambient air, led to initial concern over potential exchange with ambient air. However, we do not think that is likely as the visible scatter already appeared within one day of storage, was not directed towards isotopic signatures of ambient air and did not increase over time. The z scores show accurate values for δ^{2}H (except after 3 days) and more questionable values for δ^{18}O. The average z-score was 0.3 ± 2.7 for δ^{18}O and 1.4 ± 2.5 for δ^{2}H.

3.2 Experiment II: Memory

In this experiment, the initial standard filled into the bags was L23, followed by cycles of filling and emptying with standard H23. This standard sequence was reversed in the second part of the experiment (initially H23, then cycles of L23). No clear memory effect was found in the first part of the experiment (Fig. 4b), whereas a clear memory effect was observed after the first filling (L1) of the second part of the experiment (Fig. 4c). However, this memory almost disappeared again in the next repetition (L2).

As depicted in Fig. 4 (a and c), except for L1, almost all measurements fell within the target standard deviation for δ^{18}O, while δ^{2}H values are more scattered. The same pattern can be seen for the z-scores (Fig. 4 b and c). While almost all the z-scores are in the accurate range or in the questionable range at the threshold of the accurate range, the values of L1 are clearly outside with values in the unacceptable range.

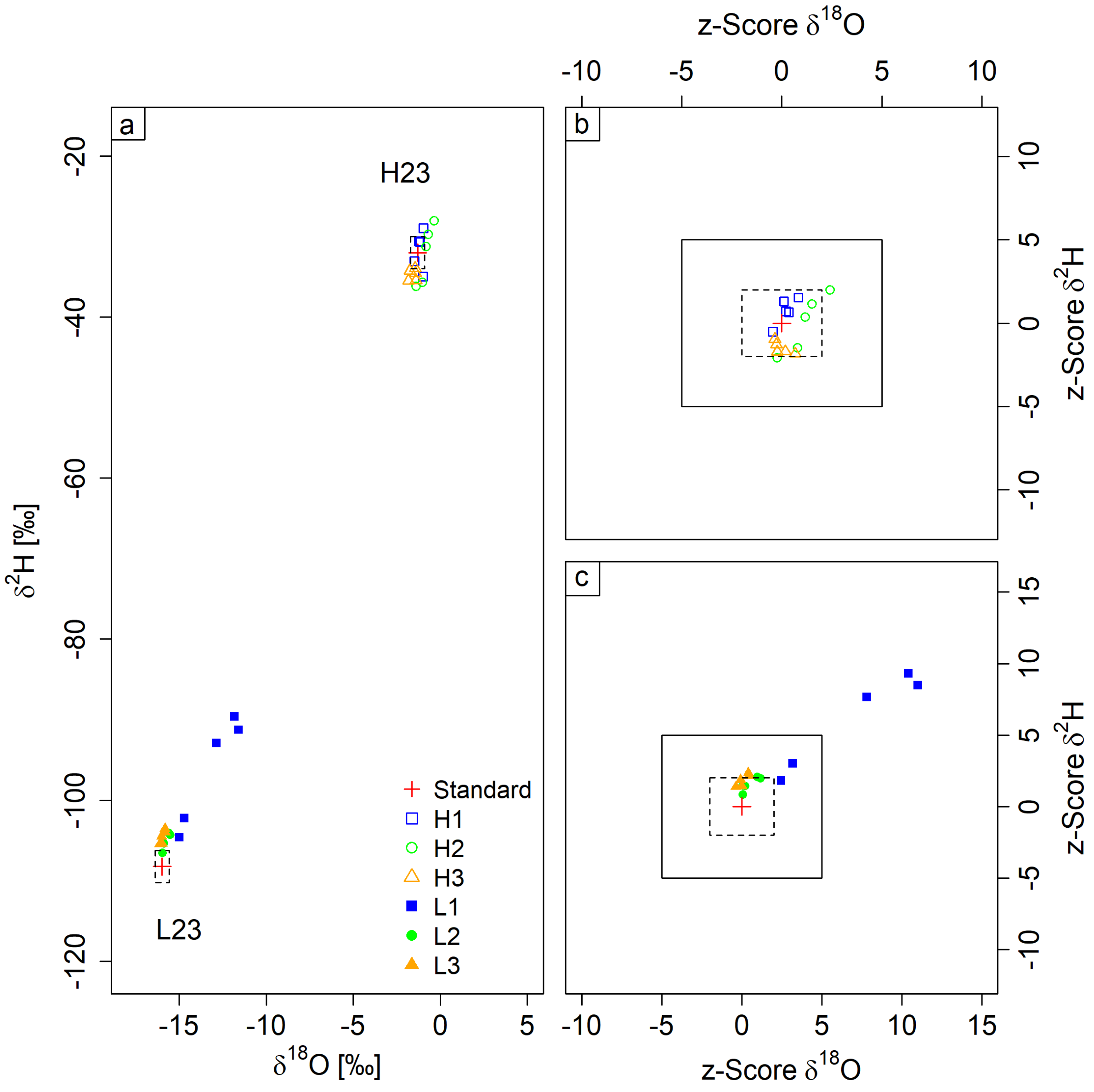

Figure 4: Memory experiment results with dual isotope plot for both experiments (a) and z-score plots for L23 to H23 (b) and H23 to L23 (c). The bags were filled first with standard H, then repeatedly (1-3) with standard L. The memory effect is evident only for measurement L1, the first to follow the change of source water vapor. The black box describes the questionable range while the scatter black box describes the accurate range (b, c). The blue dashed line (a) describes the defined standard deviation for measurements.

3.3 Experiment III: Memory test with storage

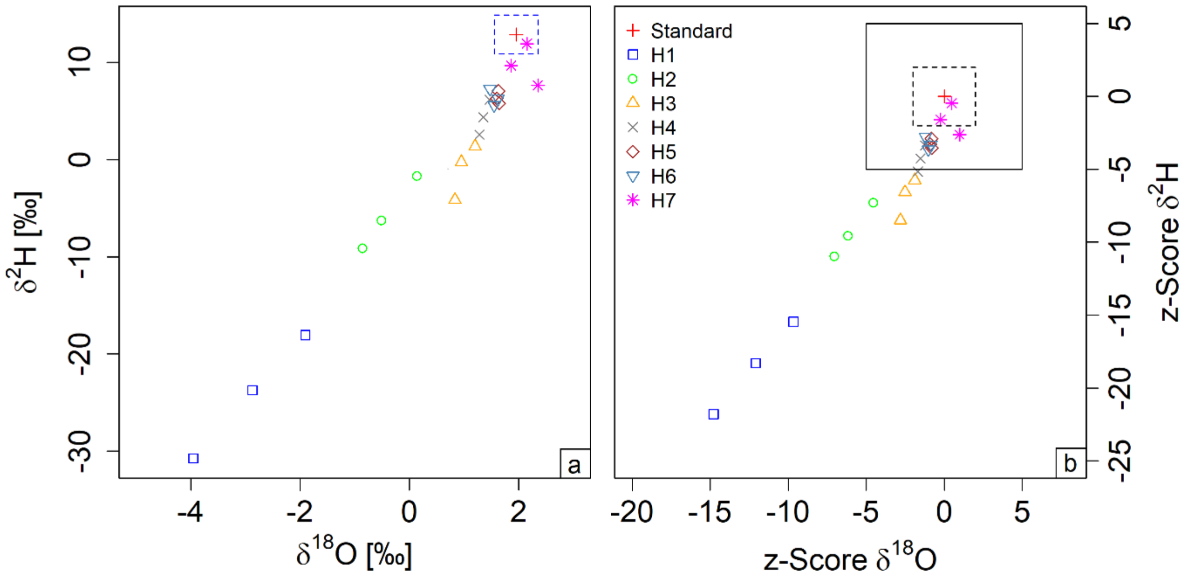

Figure 5: Memory test with storage experiment: dual isotope plot on the left and z-score plot on the right. The red cross describes the target standard value. The black box describes the questionable range while the dashed black box describes the accurate range, based on our classification of z-values (b) (Sec. 2.5). The arrow indicates the direction from strong to weak memory effect. The blue dashed line (a) describes the defined standard deviation for measurements.

No significant storage effect was observed over the one-day storage period, and there was no noticeable difference between the two repetitions (mean difference between days: 0.4 ± 0.4 ‰ $\delta^{18}O$ and 0.1 ± 1.9 ‰ $\delta^{2}H$). However, when the water source was changed to H, there was a clear memory effect of a magnitude up to -4.9 ± 1 ‰ $\delta^{18}O$ in and -37 ± 6.4 in ‰ $\delta^{2}H$ (Fig. 5 and Tab. 2). After filling with the opposite standard, H22, the first measurements (H1) revealed a low accuracy due to low precision and trueness, which was improved by around 50% with each repetition until the average result of H7 was close to the target standard value. The z-scores followed a similar trend from H1 to H5, gradually decreasing. Although H1 and H2 showed unacceptable z-scores for $\delta^{18}O$, and H3 fell within the questionable range, all subsequent measurements had z-scores within the accurate range. The $\delta^{2}H$ z-scores follow a similar trend to the z-scores for $\delta^{18}O$, thus also indicating a clear memory effect. However, this effect persisted for more cycles in the case of $\delta^{2}H$. The measurements H1 to H3 were in the unacceptable range, while the results for H4 to H6 were questionable.

Table 2: Mean differences between measured and known isotopic signatures (S-B, eq. 4) of the different repetitions of the combined storage and memory experiment.

Repetition	**Diff. $\delta^{18}O$ [‰]**	**Diff. $\delta^{2}H$ [‰]**
H1	-4.9 ± 1	-37 ± 6.4
H2	-2.4 ± 0.5	-18.6 ± 3.7
H3	-1 ± 0.2	-13.9 ± 2.8
H4	-0.6 ± 0.1	-8.5 ± 1.8
H5	-0.3 ± 0	-6.5 ± 0.7
H6	-0.4 ± 0.1	-6.5 ± 0.9
H7	0.2 ± 0.3	-3.1 ± 2.2

3.4 Experiment IV: Field filling and bag reuse

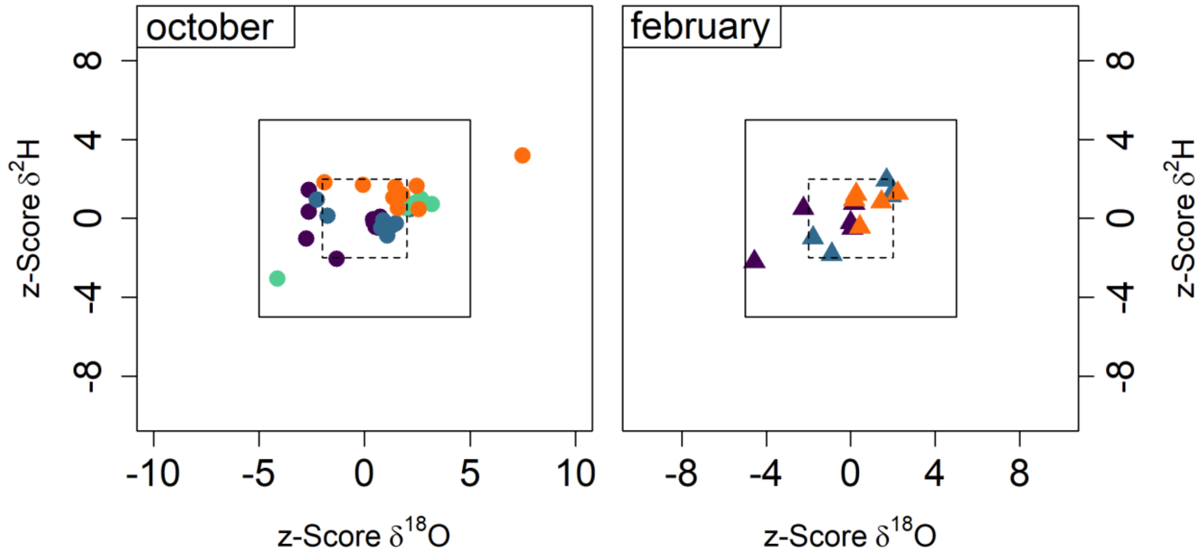

Figure 6: In October (left), in situ CRDS measurements were compared with bag measurements taken and measured directly after filling. In February (right), in situ CRDS measurements were compared with reused bags measured the next day in the laboratory (after 10 cycles of flushing with dry air).

To compare the measurements during the two campaigns and calculate the Z-scores (Eq. 4), we considered the measured isotopic value made by the CRDS in the field as the benchmark value (B) and the measurements from the gas bags as the sample (S). The average difference between direct measurement and bag measurement was 0.2 ± 0.9 ‰ for δ^{18}O and 0.7 ± 2.3 ‰ for δ^2H during the first sampling campaign in October, 2022 and 0.1 ± 0.8 ‰ for δ^{18}O and 1.4 ± 3.3 ‰ for δ^2H for the second sampling campaign with reused bags in February, 2023 (Fig. 6). The deviation of the bag method from direct in situ measurements was thus mostly within the uncertainty range of the in situ method and yielded in highly accurate z-scores for δ^2H. However, the δ^{18}O z-scores exhibit a larger scatter compared to δ^2H, consistent with the results of the laboratory storage experiment (Exp. I).

Table 3: *Mean d*ifferences between direct and bag measurement *(S-B, eq. 4)* of water stable isotopes (δ^{18}O and δ^2H) and z-scores of the different depth during the two field experiments.

Depth [cm]	**Diff. δ^{18}O [‰]**	**Diff. δ^2H [‰]**	**Z-score δ^{18}O**	**Z-score δ^2H**
		New bags		
5	- 0.3 ± 0.6	- 0.6 ± 1.9	- 0.7 ± 1.6	- 0.3 ± 1
15	0.2 ± 0.6	- 0.2 ± 1.1	0.5 ± 1.6	- 0.1 ± 0.6
45	0.6 ± 1	0.4 ± 2.9	1.4 ± 2.5	0.2 ± 1.5
150	0.8 ± 1	2.9 ± 1.6	1.9 ± 2.5	1.5 ± 0.8
		Reused bags		
5	- 0.5 ± 0.8	- 0.6 ± 2.3	- 1.3 ± 2.1	- 0.3 ± 1.2
15	0.4 ± 0.7	2.13 ± 4.2	0.9 ± 1.8	1.1 ± 2.1
150	0.4 ± 0.4	2.5 ± 2.6	1 ± 0.9	1.2 ± 1.3

3.5 Experiment V: Observation over a full cultivation period

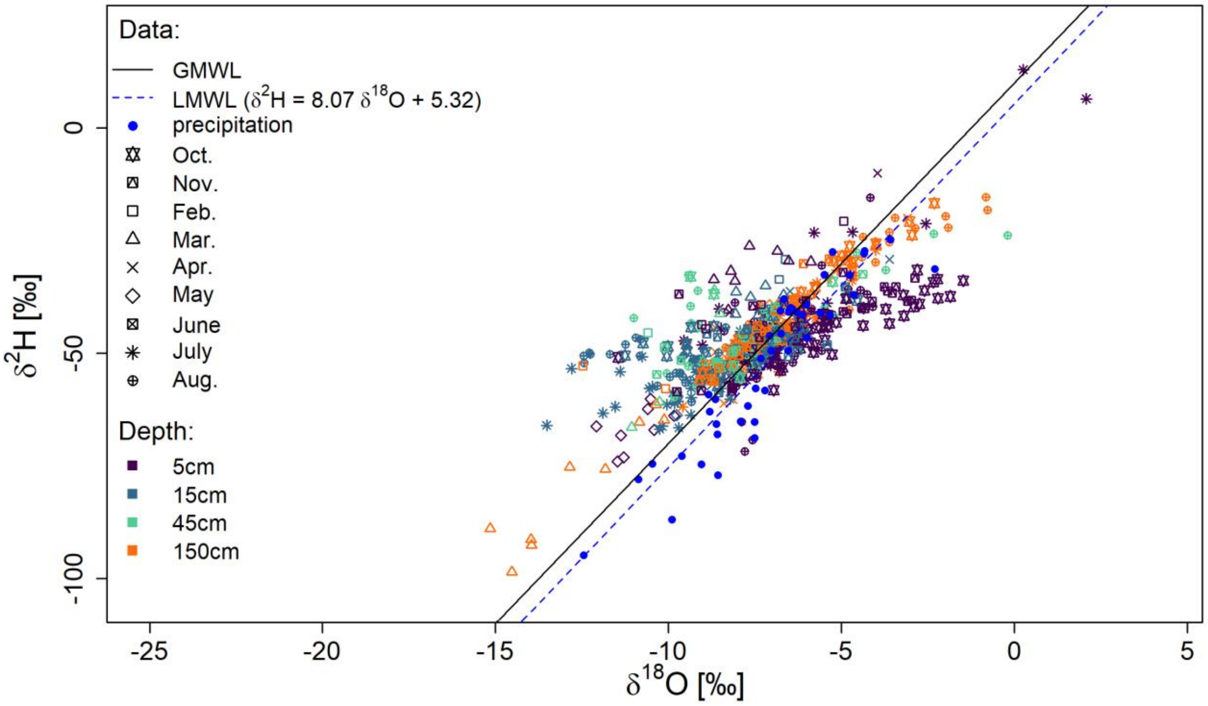

Figure 7: The dual isotope plot shows all 603 measurements taken during the cultivation period, including the Global Meteoric Water Line (GMWL; black line), the Local Meteoric Water Line (LMWL; blue dashed line) and the bag method measurements at 5, 15, 45 and 150 cm depth (purple, blue, green and yellow) during 9 different month.

Measurements of soil water isotope profiles over the full season (Fig. 7) revealed a wide range of isotopic signatures with 2.1 ‰ to -15.2 ‰ for $\delta^{18}O$ and 12.9 ‰ to -98.5 ‰ for δ^2H. Of the 623 measurements taken, 20 measurements or 3.2% had to be discarded due to damaged bags, filling errors, or condensation during the measurement and are therefore not shown (see "Handling Recommendations" in the supplement for further details). The isotopic signature of precipitation is represented by the local meteoric water line (LMWL), shown here for the period of September 2021 to September 2023. The LMWL is nearly parallel to the Global Meteoric Water Line (GMWL). In general, the measurements show isotopic signatures similar to precipitation immediately after rain events and a trend toward evaporative enrichment during droughts (see Fig. S1, supplement), but with distinct differences between months (e.g., Mar vs. Oct, at the 5 cm depth). Overall, our findings from the field trial suggest a good agreement with the LMWL and are plausible in terms of seasonal variability.

4. Discussion

Our results are generally comparable in accuracy to previous studies of water vapor storage. For example, the Soil Water Isotope Storage System (SWISS) introduced by Havranek et al. (2020) showed a high accuracy during a 30-day storage period in a laboratory experiment (± 0.5 ‰ $\delta^{18}O$ and ± 2.4 ‰ $\delta^{2}H$). This result was followed by several experiments, which showed an actual precision of 0.9 ‰ and 3.7‰ for $\delta^{18}O$ and $\delta^{2}H$ in field applications with a storage time of 14 days (Havranek et al., 2023). Their system is based on custom-made 750 ml glass vials with stainless steel connections. Magh et al. (2022) developed the vapor storage vial system (VSVS), which is based on crimp neck vials in combination with a PTFE/butyl membrane and has a similar accuracy compared to our results after one day of storage, but, like the static vials used by Havranek et al. (2020), requires a linear correction. Moreover, although the mean isotopic composition remained the same throughout the measurement, it increasingly led to high scatter of the measured isotopic signatures. Both systems are more difficult to handle during the measurement compared to inflatable bags as they must be filled with the same amount of dry gas mixture during the measurement due to the static volume of the glass vials.

A recent paper compared different types of affordable food storage bags for water vapor sampling using standardized water with different isotopic signatures (Herbstritt et al.; 2023). These authors conducted rigorous tests of diffusion tightness and inertness of various bag types. They detected significant memory in all bag types, even after flushing with dry N_2. To circumvent these memory effects, they explored preconditioning of the bags with a known moist air sample where the goal was not to eliminate the memory effect, but to make it predictable and remove it. After one storage day, the accuracy was 0.25 ± 0.41‰ and 0.41 ± 1.93 ‰ for $\delta^{18}O$ and $\delta^{2}H$. This preconditioning resembles the pre-treatment of feathers (Hobson *et al.*, 1999) and hair (Ehleringer *et al.*, 2020) to fill exchange sites with known water vapor prior to analysis, followed by post-processing to remove the pre-treatment effect.

Our study differs from Herbstritt et al. (2023) paper in several important ways. First, we have used different bags, which are more expensive, but have better control over suppliers and better description of specifications. Second, we have modified the valve inlets to the bags in a way that may reduce leakage and simplify gas transfers both in the lab and in the field. Third, we have suggested a means by which multiple flushes of the bags with dry air may eliminate, or at least minimize, the memory effect. Fourth, we have identified a time-dependent memory effect, which is consistent with the notion that some diffusion/adsorption process occurs over many hours within the walls of the bag, setting an isotopic signal that requires multiple flushes to remove. This time-dependent process does not seem to require slow flushing to reverse the memory effect (Expt. IV). These results call for the automation of bag flushing protocols in order to make these techniques routinely useful. Since the isotopic range in the experiment was relatively narrow (< 20 ‰ for δ^2H between first and second sampling), we additionally performed a small reuse experiment using two laboratory standards with higher differences in isotopic signatures and ten-times flushing with dry air (Fig. S2). As expected, results were unaffected for both δ^{18}O and δ^2H directly after bag filling. While storing did not influence the δ^{18}O signature, a clear but consistent effect was visible after one day regarding δ^2H, which, contrary to the results of Herbstritt et al. (2023), did not increase over three days storage. Since this effect was stable and we know the previous sample signature, this effect may be correctable like the moist conditioning approach described by Herbstritt et al. (2023) or erasable by increasing the number of flushes.

To the best of our knowledge, a campaign of measuring soil water isotopes using gas bags over an entire cultivation period, as shown in this study, has not been done before. However, such studies have been done with other data collection techniques. For example, the isotopic composition of water in surface soils can change significantly as evaporated soil vapor is depleted in heavy isotopes, leaving the remaining soil water enriched in ^{18}O and ^{2}H (Dubbert and Werner, 2018). This results in a wide range of isotopic signatures throughout the complete cultivation season, as can be seen in the scatter around the LMWL (Fig. 7). As expected, evaporative enrichment is evident following precipitation free periods in the upper 5 cm depth (e.g. April period in Fig. S1), but not after the rainy winter period. In contrast, there are only slight trends in evaporative enrichment at lower depths (compare e.g. Sprenger et al., 2016).

In the past, destructive measurements of soil water have relied predominantly on cryogenic vacuum extraction (CVE). The accuracy of CVE can vary greatly for soil samples and is associated with co-extraction of organic compounds, significantly interfering with the isotopic quantification using CRDS (Orlowski et al., 2016b). In comparison, methods using in situ soil or xylem probes based on semi permeable tubes have been reported to be highly accurate but complex to handle and set up (Volkmann and Weiler, 2014; Volkmann et al., 2016; Rothfuss et al., 2013; Kübert et al., 2020). Therefore, development efforts to combine destructive with in-situ sampling continue.

Recent studies showed that sampling of water vapor with subsequent analysis in the laboratory is possible with both glass bottles and different types of bags. Glass containers revealed the advantage of less material effects and higher diffusion tightness while gas bags were easier to measure due to their flexible structure. Nevertheless, further experiments should investigate the detected interaction of water samples within the gas bag wall. The cost for the commercially available gas bags we used was relatively low. To classify the costs per sample container, the SWISS-System was clearly more expensive while the other methods were less expensive per sample container with 1-2€ but produced running cost (Magh et. al, 2022) or additional cost and effort to attach the valve and built the final bag (Herbstritt et. al, 2023). We have demonstrated that commercially available our bags met the expected level of performance already, provided that they were flushed multiple times between uses, they are reused for the same sample probes, and that standards are taken through the whole sample collection, transport, and analysis process. Following the described conditions, we were able to reliably measure soil water over a full cultivation period under natural abundance conditions.

4. Conclusion

Our laboratory and field experiments have confirmed that soil membranes combined with gas bags for in situ soil water vapor sampling and subsequent water stable isotope analyses was a reliable, cost-effective, and easy to handle method allowing for many future applications. We were able to demonstrate that both 1) storage is possible and 2) memory effects caused by previous samples can be prevented by appropriate preconditioning, allowing the gas bags to be reused. Most importantly, we recommend rapid dry-air flushing rather than the slow pre-conditioning with humid air recommended by Herbstritt et al. (2023). Regarding the isotopic signature during the experiment, reuse is easier to carry out with smaller differences between the consecutive samples in the bags, e.g. in the natural abundance range. However, for larger differences in isotopic signatures, the bags need to be handled differently (e.g. better flushing between samples or no reuse). Through the conducted field experiment (two campaigns with CRDS and bag measurements), we were able to show that the bags could be used in our case with an accuracy of 0.23 ± 0.84 δ^{18}O [‰] and 0.94 ± 2.69 δ^{2}H [‰], which allows a wide applicability. The possibility to take and store samples easily and without permanent power supply extends the usability of water stable isotope measurements in the field.

5. Data availability

The data will be available in the BonaRes repository upon publication.

6. Author contribution

AD and MD designed the study. AD conducted the experiments and analyzed the data. JM, DD, and MH provided support for the experimental setup and analysis methods. AD prepared the paper with supervision from MD and contributions from all co-authors.

7. Competing interests

The authors declare that they have no conflict of interest.

8. Acknowledgements

We acknowledge funding by ZALF Leibniz as well as the Leibniz association (ISO-SCALE project; project number K444/2022). The authors wish to thank the Experimental Infrastructure Platform (EIP) of ZALF and Linda Röderer for assisting with the field experiments.

9. References

Craig, H.: Standard for Reporting Concentrations of Deuterium and Oxygen-18 in Natural Waters, Science, 133, 1833–1834, https://doi.org/10.1126/science.133.3467.1833, 1961.

Dubbert, M., and Werner, C.: Water fluxes mediated by vegetation: emerging isotopic insights at the soil and atmosphere interfaces. New Phytol, 221(4), 1754-1763, https://doi.org/10.1111/nph.15547, 2018.

Ehleringer, J. R., Bowen, G. J., Chesson, L. A., West, A. G., Podlesak, D. W., & Cerling, T. E.: Hydrogen and oxygen isotope ratios in human hair are related to geography. P Natl Acad Sci-Biol, 105(8), 2788-279, https://doi.org/10.1073/pnas.0712228105, 2008.

Gat, J. R.: Oxygen and hydrogen isotopes in the hydrologic cycle. Ann Rev Earth Pl Sc, 24(1), 225-262, https://doi.org/10.1146/annurev.earth.24.1.225, 1996.

Havranek, R. E., Snell, K. E., Davidheiser-Kroll, B., Bowen, G. J., and Vaughn, B.: The Soil Water Isotope Storage System (SWISS): An integrated soil water vapor sampling and multiport storage system for stable isotope geochemistry. Rapid Commun Mass Sp, 34(12), e8783, https://doi.org/10.1002/rcm.878, 2020.

Havranek, R. E., Snell, K., Kopf, S., Davidheiser-Kroll, B., Morris, V., and Vaughn, B.: Lessons from and best practices for the deployment of the Soil Water Isotope Storage System. Hydrol Earth Syst Sc, 27(15), 2951-2971, https://doi.org/10.5194/hess-27-2951-2023, 2023.

Herbstritt, B., Gralher, B., Seeger, S., Rinderer, M., and Weiler, M.: Discrete in situ vapor sampling for subsequent lab-based water stable isotope analysis. Hydrol Earth Syst Sc, 27(20), 3701-3718, https://doi.org/10.5194/hess-27-3701-2023, 2023.

Hobson, K. A.: Tracing origins and migration of wildlife using stable isotopes: a review. Oecologia, 120, 314-326, https://doi.org/10.1007/s004420050865, 1999.

Horita, J., and Wesolowski, D. J.: Liquid-vapor fractionation of oxygen and hydrogen isotopes of water from the freezing to the critical temperature. Geochim Cosmochim Ac, 58(16), 3425-3437, https://doi.org/10.1016/0016-7037(94)90096-5, 1994.

Kübert, A., Paulus, S., Dahlmann, A., Werner, C., Rothfuss, Y., Orlowski, N., and Dubbert, M.: Water stable isotopes in ecohydrological field research: comparison between in situ and destructive monitoring methods to determine soil water isotopic signatures. Front plant sci, 11, 497124, https://doi.org/10.3389/fpls.2020.00387, 2020.

Kühnhammer, K., Dahlmann, A., Iraheta, A., Gerchow, M., Birkel, C., Marshall, J. D., and Beyer, M.: Continuous in situ measurements of water stable isotopes in soils, tree trunk and root xylem: Field approval. Rapid Commun Mass Sp, 36(5), e9232, https://doi.org/10.1002/rcm.9232, 2021.

Kühnhammer, K., van Haren, J., Kübert, A., Bailey, K., Dubbert, M., Hu, J., Ladd, N. S., Meridith, L. K., Werner, C., and Beyer, M.: Deep roots mitigate drought impacts on tropical trees despite limited quantitative contribution to transpiration. Sci Total Environ, 893, 164763, https://doi.org/10.1016/j.scitotenv.2023.164763, 2023.

Marshall, J.D., Cuntz, M., Beyer, M., Dubbert, M., Kuehnhammer, K.: Borehole equilibration: Testing a new method to monitor the isotopic composition of tree xylem water in situ. Front Plant Sci., 11, 508657, https://doi.org/10.3389/fpls.2020.00358, 2020.

Magh, R. K., Gralher, B., Herbstritt, B., Kübert, A., Lim, H., Lundmark, T., and Marshall, J.: Conservative storage of water vapour–practical in situ sampling of stable isotopes in tree stems. Hydrol Earth Syst Sc, 26(13), 3573-3587, https://doi.org/10.5194/hess-26-3573-2022, 2022.

Majoube M. Fractionnement en oxygene 18 et en deuterium entre l'eau et sa vapeur. J Chim Phys, 68, 1423-1436, https://doi.org/10.1051/jcp/1971681423, 1971.

Mook, W. G. (Eds.): Environmental isotopes in the hydrological cycle: principles and applications. UNESCO/IAEA, Volume 1, Centre for Isotope Research, Groningen, Netherlands, 280 pp., 2000.

Orlowski N., Pratt D.L., and McDonnell J.J.: Intercomparison of soil pore water extraction methods for stable isotope analysis. Hydrol Process, 30(19), 3434-3449, https://doi.org/10.1002/hyp.10870, 2016a.

Orlowski N., Breuer L., Mcdonnell J.J.: Critical issues with cryogenic extraction of soil water for stable isotope analysis. Ecohydrology, 9(1), 3-10, https://doi.org/10.1002/eco.1722, 2016b.

Rothfuss, Y., Biron, P., Braud, I., Canale, L., Durand, J. L., Gaudet, J. P., Richard P., Vauclin M., and Bariac, T.: Partitioning evapotranspiration fluxes into soil evaporation and plant transpiration using water stable isotopes under controlled conditions. Hydrol Process, 24(22), 3177-3194, https://doi.org/10.1002/hyp.7743, 2010.

Rothfuss, Y., Vereecken, H., and Brüggemann, N.: Monitoring water stable isotopic composition in soils using gas-permeable tubing and infrared laser absorption spectroscopy. Water resour res, 49(6), 3747-3755, https://doi.org/10.1002/wrcr.20311, 2013.

Sense Trading B.V., personal communication, 2024.

Séraphin, P., Vallet-Coulomb, C., and Gonçalvès, J.: Partitioning groundwater recharge between rainfall infiltration and irrigation return flow using stable isotopes: The Crau aquifer. J Hydrol, 542, 241-253, https://doi.org/10.1016/j.jhydrol.2016.09.005, 2016.

Sodemann, H. (Eds.): Tropospheric transport of water vapour: Lagrangian and Eulerian perspectives. Swiss. ETH Zurich, No. 16623, 225 pp., ISBN 978-3832513849, 2006.

Sprenger M, Herbstritt B, Weiler M. Established methods and new opportunities for pore water stable isotope analysis. Hydrol Process. 29(25), 5174-5192, https://doi.org/10.1002/hyp.10643, 2015.

Sprenger, M., Leistert, H., Gimbel, K., and Weiler, M.: Illuminating hydrological processes at the soil-vegetation-atmosphere interface with water stable isotopes. Rev Geophys, 54(3), 674-704, https://doi.org/10.1002/2015RG000515, 2016.

Volkmann, T. H., and Weiler, M.: Continual in situ monitoring of pore water stable isotopes in the subsurface. Hydrol Earth Syst Sc, 18(5), 1819-1833, https://doi.org/10.5194/hess-18-1819-2014, 2014.

Volkmann, T. H., Haberer, K., Gessler, A., and Weiler, M.: High-resolution isotope measurements resolve rapid ecohydrological dynamics at the soil–plant interface. New Phytol, 210(3), 839-849, https://doi.org/10.1111/nph.13868, 2016.

Wassenaar, L. I., Ahmad, M., Aggarwal, P., Van Duren, M., Pöltenstein, L., Araguas, L., and Kurttas, T.: Worldwide proficiency test for routine analysis of δ^{2}H and δ^{18}O in water by isotope-ratio mass spectrometry and laser absorption spectroscopy. Rapid Commun Mass Sp, 26(15), 1641-1648, https://doi.org/10.1002/rcm.6270, 2012.

West, A. G., Patrickson, S. J., and Ehleringer, J. R.: Water extraction times for plant and soil materials used in stable isotope analysis. Rapid Commun Mass Sp, 20(8), 1317-1321, https://doi.org/10.1002/rcm.2456, 2006.