



HAL
open science

Improving memory effect correction to achieve high precision analysis of $\delta^{17}\text{O}$, $\delta^{18}\text{O}$ $\delta^2\text{H}$, ^{17}O -excess and d-excess in water using cavity ring-down laser spectroscopy

Christine Vallet-Coulomb, Martine Couapel, Corinne Sonzogni

► To cite this version:

Christine Vallet-Coulomb, Martine Couapel, Corinne Sonzogni. Improving memory effect correction to achieve high precision analysis of $\delta^{17}\text{O}$, $\delta^{18}\text{O}$ $\delta^2\text{H}$, ^{17}O -excess and d-excess in water using cavity ring-down laser spectroscopy. *Rapid Communications in Mass Spectrometry*, 2021, 10.1002/rcm.9108 . hal-03209687

HAL Id: hal-03209687

<https://hal.science/hal-03209687>

Submitted on 20 Feb 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Improving memory effect correction to achieve high precision analysis of $\delta^{17}\text{O}$, $\delta^{18}\text{O}$**
2 **$\delta^2\text{H}$, ^{17}O -excess and d-excess in water using cavity ring-down laser spectroscopy**

3

4 Christine Vallet-Coulomb¹, Martine Couapel¹, Corinne Sonzogni¹

5 ¹ Aix Marseille Univ, CNRS, IRD, INRA, Coll France, CEREGE, Aix-en-Provence, France

6

7

8 **Abstract**

9 **Rationale:** The precision obtained in routine isotope analysis of water ($\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$, ^{17}O -
10 excess and d-excess) by Cavity Ring-Down Spectroscopy is usually below the instrument
11 specifications provided by the manufacturer. This study aimed at reducing this discrepancy,
12 with particular attention paid to mitigating the memory effect (ME).

13 **Methods:** We used a Picarro L2140i analyzer coupled with a high-precision A0211 vaporizer
14 and an A0325 autosampler. The magnitude and duration of the ME were estimated using 24
15 series of 50 successive injections of samples with contrasting compositions. Four memory
16 correction methods were compared, and the instrument performance was evaluated over a 17-
17 month period of routine analysis, using two different run architectures.

18 **Results:** The ME remains detectable after the 30th injection, implying that common corrections
19 procedures only based on the last preceding sample need to be revised. We developed a new
20 ME correction based on the composition of several successive samples, and designed a run
21 architecture to minimize the magnitude of the ME. The standard deviation obtained from
22 routine measurement of a QA water over a 7 months-period was 0.015‰ for $\delta^{17}\text{O}$, 0.023‰ for
23 $\delta^{18}\text{O}$, 0.078‰ for $\delta^2\text{H}$, 0.006‰ for ^{17}O -excess and 0.173‰ for d-excess. In addition, we
24 provided the first $\delta^{17}\text{O}$ and ^{17}O -excess values for the GRESP certified reference material.

25 **Conclusions:** This study demonstrates the long-term persistence of the ME, which is often
26 overlooked in routine analysis of natural samples. As already evidenced when measuring
27 labelled water, it calls for consideration of the compositions of several previous samples to
28 obtain an appropriate correction, a prerequisite to achieve high precision data.

29

30 1- Introduction

31
32 Laser absorption spectrometric techniques are more accessible and easier to operate than
33 Isotope Ratio Mass Spectrometers (IRMS), which has led to a growing interest in producing
34 water isotope data, as revealed by the increasing number of laboratories involved in the
35 international proficiency tests organized by the International Atomic Energy Agency (IAEA)
36 between 2011 and 2016.^{1,2} The most recent instruments offer the opportunity to measure the
37 $^{17}\text{O}/^{16}\text{O}$ ratio of water in addition to the traditionally analyzed ratios of $^{18}\text{O}/^{16}\text{O}$ and $^2\text{H}/^1\text{H}$. The
38 analysis of all three oxygen isotopes allows one to calculate the ^{17}O -excess value, defined as
39 $^{17}\text{O}\text{-excess} = \delta^{17}\text{O} - 0.528 \times \delta^{18}\text{O}$ (with $\delta' = \ln(\delta+1)$; all values in the decimal notation),
40 similarly to the classic d-excess value ($\text{d-excess} = \delta^2\text{H} - 8 \times \delta^{18}\text{O}$). Regarding the natural
41 variability of the ^{17}O -excess in meteoric waters,³ a measurement precision better than 10^{-5} (10
42 permeg) is required to allow the application of this new tracer in water cycle studies. This
43 remains challenging in routine analysis, and appropriate data processing and measurement
44 protocols are necessary to resolve the small variations of ^{17}O -excess observed in nature. Post
45 measurement data processing involves compensation for instrument drift and memory effect,
46 and VSMOW-SLAP normalization.⁴⁻⁷ Additionally, regular measurements of a quality control
47 standard are required to ensure long-term reproducibility with high precision.

48
49 The so-called “memory effect” (henceforth ME), affecting isotopic measurements due to the
50 carry-over of residual amounts of water during sample analysis, is an important pitfall that
51 limits the precision obtained when measuring disparate samples under routine conditions.^{5,8-10}
52 For the Picarro wavelength-scanned cavity ring-down spectrometers (WS-CRDS, Picarro, Santa
53 Clara, CA, USA), the manufacturer specifies the magnitude of the ME, as a percentage of the
54 final value, being better than 98% (for $\delta^2\text{H}$) and 99% (for $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$) after the 4th injection,
55 and recommends to simply discard the first few injections. Nevertheless, the ME remaining
56 after the 4th injection may strongly affect measurement precision, especially when measuring
57 samples with significant differences in isotopic composition, as the reference waters used for
58 calibration, or even more when analyzing strongly enriched (labelled) waters. As an illustration,
59 for a difference in $\delta^2\text{H}$ composition of ± 60 ‰ between consecutive samples, which is a typical
60 value encountered when measuring batches of precipitation samples, the ME would induce an
61 offset of ± 1.2 ‰ from the true value. Therefore, except in cases when innovative modifications
62 of the internal measurement settings could be developed to mitigate the ME,⁸ it remains highly

63 recommended to systematically apply a numerical correction of the ME during post-
64 measurement data processing.

65
66 In usual practices of laser data processing, the magnitude of the ME is generally evaluated
67 during each run using reference waters included in the measurement batches, and corrected in
68 consideration of the difference in isotopic composition between 2 successive samples.^{4-7,9,11} In
69 most cases, satisfying precisions can be obtained. Nevertheless, the succession of injections is
70 similar when the same sample is injected, and when moving to the following sample, except
71 for the syringe rinses performed between vials. Therefore, assuming that the measurement of a
72 sample is only affected by the last preceding one would imply that the persistence of the ME is
73 equivalent to the time (or injection number) between two samples. The number of successive
74 injections of the same sample hardly overtake 6-8 during routine analysis (or even less) to
75 maintain reasonable data throughput.^{4-6,9,11,12} An overlooked longer persistence of memory
76 would challenge the current practices, since more than one preceding sample could be involved
77 in the ME.

78
79 It is generally assumed that 10-15 injections of the same sample are sufficient to obtain a
80 composition free of memory.^{4-6,9,10} While the exact duration of the ME has never been fully
81 assessed, some previous studies already suggested that substantial ME could remain after 15
82 injections.⁷ In addition, no evaluation has been performed so far of the impact of the ME on
83 ¹⁷O-excess data, despite the high level of precision required to obtain reliable values. As the
84 ME is expected to be the main factor reducing the precision and accuracy compared to the basic
85 specifications of the instrument,⁸ an in-depth exploration of its impact and of the appropriate
86 correction strategies is required.

87
88 The objectives of the present study were therefore to: 1) estimate the magnitude, duration, and
89 variability of the ME and its impact on data precision and accuracy ; 2) propose a new correction
90 method for the ME during post-run data processing and compare it with current practices,
91 including the multi-reservoir correction method which has been developed for measuring
92 labelled (*i.e.* strongly enriched) water¹³; 3) propose an optimal architecture for routine runs, and
93 4) evaluate the resulting improvement in measurement precision for the stable oxygen and
94 hydrogen isotopes of the water molecule ($\delta^{17}\text{O}$, $\delta^{18}\text{O}$ and $\delta^2\text{H}$) and their derived tracers (¹⁷O-
95 excess and d-excess).

96

97 2- Method

98

99 2-1 Basic protocol and routine measurement runs

100

101 We used a commercially available Picarro L2140i isotopic analyzer (Picarro, Santa Clara, CA,
102 USA) to simultaneously measure $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ and derive the values of ^{17}O -excess and d-
103 excess. Water samples were stored in 2-ml glass vials and injected with a liquid autosampler
104 (Picarro A0325) into a vaporizer module (Picarro A0211), using pure nitrogen as the carrier
105 gas. A 10- μL syringe was used to collect the samples. The syringe was rinsed between each
106 vial, and purged twice with the sample water, prior to injection into the vaporizer. The injection
107 volume was 1.6 μL , the injection speed was 1 $\mu\text{L} \times \text{s}^{-1}$, and throughout the study, only the ^{17}O -
108 mode of the analyzer was used. The initial instrument specifications for liquid water
109 measurement provided by the manufacturer, estimated as the standard deviation (SD) of the
110 mean of groups of 6 injections over 8 hours, were 0.025‰, 0.025‰, 0.1‰ and 0.015‰ for
111 $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ and ^{17}O -excess, respectively.

112

113 Two different run architectures were used for routine analysis (Table 1). The run architecture
114 #1, used until June 2019, was designed to have an average duration of 48h, *i.e.* a total of about
115 320 injections, divided into 8 injections of 40 water samples. The run included 2 to 3 replicates
116 of all samples in order to obtain the best precision possible for all isotopes and their derived
117 secondary parameters ^{17}O -excess and d-excess. Normalization to the VSMOW-SLAP scale was
118 performed by using a set of 3 in-house Working Standards (WS in the following) measured at
119 the beginning, middle, and end of each run in order to monitor instrument drift.⁵ A linear
120 calibration was applied, and the linearity of the calibration was evaluated by comparing the
121 scaling factors corresponding for the different intervals between WS. These in-house WS were
122 named ICE (low δ value), SEA (high δ value) and TAP (intermediate composition). A fourth
123 in-house standard (ROB) was dedicated to Quality Assurance (QA), and treated as an unknown
124 sample in all runs (the same number of replicates as for the samples). To normalize the $\delta^{17}\text{O}$
125 value in the VSMOW-SLAP scale, we used the SLAP2 composition proposed by Schoenemann
126 et al. (2013).¹⁴

127 From September 2019, a different run architecture was adopted (run architecture #2). The
128 modification was based on the evidence of a ME higher than expected, and after having tested
129 different numbers of injections (12 to 15 injections, applied either to all vials or to WS only).
130 In the new scheme, two sets of WS were measured, at the beginning and at the end of a run. A

131 “symmetrical ordering” was adopted for the two sets of WS (ascending *versus* descending delta-
132 values), and for sample replicates. The number of injections was increased to 15 for calibration
133 WS, while keeping 8 injections per sample. This organization was expected to mitigate most of
134 the ME when averaging replicates, especially for intermediate composition ranges. However,
135 to tackle the problem of very low and very high isotopic compositions, each group of calibration
136 WS was preceded by a conditioning vial filled with the same water (Table 1). The data
137 throughput was almost the same for both versions of the routine run architecture, with a total
138 of 30 unknown water samples measured during approximately 48 hours, *i.e.* 15 replicated
139 samples including the QA samples.

140

141 The post measurement data processing was carried out in the Python programming language¹⁵
142 (script available in supplementary information). All sample compositions are provided on the
143 VSMOW-SLAP scale, after ME correction. In addition, the following information was
144 systematically gathered in a “run summary” file to verify the overall quality of each
145 measurement run: raw composition of the WS, mean water concentration and associated
146 standard deviation for all injections of the run, calibrated values of the QA standard. Plots are
147 also drawn to visualize all raw values and water concentration during the run, and a focus on
148 all injections of the WS.

149

150 **2-2 Measurement run dedicated to assessing the memory effect**

151

152 The magnitude of the ME can be estimated from a memory factor representing the proportion
153 of residual water vapor in the instrument during a measurement:

154

$$155 \quad x_i = \frac{\delta_{i,meas} - \delta^n}{\delta^r - \delta^n} \quad \text{Eq. (1)}$$

156

157 where, during the injection i of a sample n (composition δ^n), x_i is the proportion of residual
158 vapor (composition δ^r), and $\delta_{i,meas}$ is the measured composition. Assuming that the
159 composition of the residual vapor corresponds to that of the previous sample (*i.e.* $\delta^r = \delta^{n-1}$),
160 x_i is equivalent to previous formulations proposed to quantify the deviation of a measurement
161 from its expected value (*e.g.* the term “ m ” in Guidotti et al.(2013),¹³ Van Geldern et al. (2011),⁹
162 or “ Φ ” in Olsen et al. (2006)¹⁶). Here it should be mentioned that the term “residual vapor”
163 will be used in the following since all measurements are in the vapor phase, but strictly

164 speaking, the ME could also result from water adsorbed in its liquid form on the different parts
165 of the measurement line, and thereafter desorbed.

166

167 The assumption that x_i represents the entire magnitude of the ME is only verified if the
168 compositions of the two successive samples are measured completely free of ME. A run
169 dedicated to evaluate the ME factors was designed with a total of 50 injections per sample vials.
170 This number was chosen to keep a reasonable duration of the run (3 days), and represents a
171 compromise between obtaining memory free values, and the effect of a potential drift. To
172 encompass all the differences in δ values between the WS, the sample vials were sorted in both
173 ascending and descending isotopic composition. In addition, three successions of vials filled
174 with identical water were included to evaluate the robustness of the number of injections. Two
175 similar dedicated runs were performed in October and November 2019. In addition, 3 other
176 runs including 50 injections of successive samples were included in the ME analysis.

177

178 **3- Results**

179

180 **3-1- Calibration of working standards and QA water**

181

182 The compositions of the three in-house WS used for calibration were determined in two runs
183 that included the VSMOW2-VSLAP2 international standards for calibration and
184 normalization,^{14,17} and the GRESP certified reference water¹⁸ for quality control (Table 2).
185 Sample vials were sorted, first, in a decreasing order of δ values, and then replicated in the
186 reverse order, such that the symmetric succession of water compositions compensated for the
187 possible ME. In addition, conditioning vials were added for VSMOW and SLAP. For each vial
188 50 successive injections were performed, and the corresponding selected raw value was
189 estimated as the average of the last 15 injections. The total run lasted 4 days (715 injections)
190 with a vaporizer septum change in the middle. A second run with the same architecture was
191 performed one week apart. The normalized isotopic compositions of the 3 WS and of the
192 GRESP standard were calculated from the average calibrated values of each of the 4 half-
193 sequences. The composition of the 3 WS used for calibration covers approximately the
194 VSMOW-GRESP range of isotopic composition. The measured $\delta^{18}\text{O}$ and $\delta^2\text{H}$ values for the
195 GRESP water fell in the range of the total combined uncertainties of IAEA composition.¹⁸ In
196 addition, we determined the first values of $\delta^{17}\text{O}$ and ^{17}O -excess for the GRESP certified
197 reference material:¹⁸ $\delta^{17}\text{O} = -17.7784 \pm 0.02 \text{ ‰}$; ^{17}O -excess = $0.025 \pm 0.005 \text{ ‰}$ (n=4).

198 Regarding the significantly greater precision obtained for ^{17}O -excess compared to $\delta^{17}\text{O}$ and
199 $\delta^{18}\text{O}$, and following the recommendations of Schoenemann et al.,¹⁴ the $\delta^{17}\text{O}$ value and $\delta^{18}\text{O}$ are
200 reported here and in Table 2 with a number of decimals higher than allowed by their absolute
201 precisions, in order to be consistent with the ^{17}O -excess calculation.

202

203 The assigned composition of the in-house standard used for QA (ROB) was determined from
204 four dedicated runs, including the SEA and TAP in-house standards : two runs with 25
205 injections of all vials, taking the last 5 injections, and two runs with 50 injections, taking the
206 last 15 injections (Table 2). The results of both runs were in good agreement, comprising
207 differences less than 0.001, 0.002, 0.022 ‰ for $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, and $\delta^2\text{H}$, respectively.

208

209 **3-2- Memory effect**

210

211 **3-2-1 Evaluation of memory effect factors**

212

213 Regarding the expected asymptotic shape of the ME, obtaining “memory-free” values needs to
214 be considered with regards to the measurement precision. The two runs dedicated to assess the
215 ME (cf. section 2-2) indicated that the ME was no longer detectable after injection #45 of the
216 same sample: the SD calculated for sets of similar water samples were of the order of the
217 instrument precision (Table 3). Nevertheless, a closer look at the succession of data pointed that
218 for $\delta^2\text{H}$ values SEA was lowest after having followed ICE, while ICE was highest after having
219 just followed SEA, suggesting that a slight ME remained after 45 injections. As stated above,
220 the total of 50 injections represents a compromise between memory removing and the effect of
221 a potential drift, regarding the total duration of the dedicated runs (>3days), which is consistent
222 with the measurement precision.

223

224 The last 5 injections of each sample were then used as the reference compositions for the current
225 sample ($\delta^n = \delta_{i,meas}$ in Eq. 1) and for the residual vapor ($\delta^r = \delta^{n-1}$). In the run architecture,
226 the succession of water sample compositions covered 3 different isotopic differences, in both
227 directions, from which 6 series of ME factors (x_i) were calculated following Equation 1. In
228 addition to these two runs, we also calculated ME factors from 3 other runs that involved 50
229 injections of successive samples with significant isotopic differences. Two of these runs
230 (February 2020) included the VSMOW2 and SLAP2 standards, and therefore covered the

231 largest differences in δ -values. This provided a total of 24 series of ME factors. The averages
232 and corresponding standard deviations are shown in Figure 1A and 1B. The highest variability
233 of ME factors was found for $\delta^{17}\text{O}$, and the lowest for $\delta^2\text{H}$ (Figure 1G). This is consistent with
234 the relative differences in the signal/noise ratio for each of the isotopic compositions, and
235 roughly corresponds to their respective precisions compared to their natural variation ranges.
236 For each isotopic species, the variability was especially high for the first 2 injections. Potential
237 drivers of variations in ME factors were tested but did not evidence any clear trend. No relation
238 was found for ME factors with the magnitude or the direction of isotopic differences between
239 samples. In addition, the time variation of the ME was evaluated by plotting the average x_i
240 values of each run against time for the injections #3 to #8 (Figure 2). This plot shows the
241 decreasing ME with more injections, as already illustrated in Figure 1, but no variation with
242 time appeared with regards to the SD of each ME factor average value. No change in the
243 measurement procedure that might explain a variation in the ME (*e.g.* a change in the injection
244 method, in the injection volume, the analyze of saline water, or a vaporizer cleaning) was
245 performed. In the absence of a clear trend, the average series of ME factors were considered to
246 be stable.

247

248 A semi-logarithmic scale allowed to evidence the exponential behavior of the ME decay with
249 time, which can be fitted with a 2-term exponential function (Figure 1C and 1D). Similar
250 average ME factors were found for $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$, showing a similar behavior for the two
251 oxygen isotopic species with regards to memory (Figure 1A and 1C). For $\delta^2\text{H}$ (Figure 1B and
252 1D), the ME was stronger and persisted longer than for $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ (Figure 1A and 1C). The
253 absolute impact of the ME with respect to the instrument precision is illustrated for the
254 transition between the high- and the low-value WS (Figure 1E and 1F), showing the need for a
255 correction until 36 injections for $\delta^{17}\text{O}/\delta^{18}\text{O}$, and 41 injections for $\delta^2\text{H}$. As the injections are
256 fully automated, the number of injections can be translated into a time scale, providing that the
257 same injection protocol is maintained: with our measurement conditions the ME needs
258 approximately 5 hours to disappear. It should be specified that some exploratory experiments
259 suggested that a lower injection volume, and thus a lower vapor concentration during pulse
260 analysis, would increase the magnitude and the duration of the ME. Further data would be
261 necessary to quantify this trend.

262

263 Reference series of ME factors were established using a smoothing of the average ME factors
264 calculated from the 24 individual ME factor series, assuming a progressive decrease in ME
265 from one injection to the next until its complete disappearance. Smoothing was performed with
266 the Python programming language¹⁵ using Sequential Least Squares Programming (SLSQP in
267 the SciPY module of PYTHON). The minimization of the difference between smoothed and
268 raw ME factors was subjected to two constraints: 1) a decreasing trend, and 2) null values for
269 the last 5 injections, used to estimate the reference composition. Smoothing was performed to
270 avoid adding some noise to the data when applying an ME correction. The smoothing is based
271 on the method proposed by Van Geldern and Barth⁹ which aimed at reducing the variance of
272 all corrected injections for a vial, after the application of an increasing correction coefficient.
273 Note that an alternative would be to use the exponential fits (Figure 1C and 1D), but the
274 smoothing better fitted the measured ME factors. Two distinct reference series were established
275 for $\delta^{18}\text{O}$ and $\delta^2\text{H}$, respectively. Based on the similarity between ME factors obtained for the
276 two oxygen isotopic species (Figure 1A), the reference series built for $\delta^{18}\text{O}$ was also used to
277 correct $\delta^{17}\text{O}$ measurements. Therefore, the application of the ME correction did not affect the
278 precision of ^{17}O -excess.

279

280 **3-2-2 Memory effect correction**

281

282 A mathematical description of the ME is needed to apply a correction. Mixing equations are
283 commonly used, but they may correspond to two different conceptual frameworks. In the first
284 approach, the composition of the residual vapor is taken as that of the preceding sample. The
285 contribution of the residual vapor decreases with successive injections, as residual vapor is
286 flushed, but the end-member compositions are assumed to be constant. Most of the classical
287 methods used to correct for the ME are based on this conceptual scheme, implying that $\delta^r =$
288 δ^{n-1} in Equation 1.^{4-7,9,11,19} In the second approach, the residual vapor exchanges constantly
289 with the current sample, and its composition progressively evolves towards that of the current
290 sample, while the exchange rate is kept constant. This is the basis of the sophisticated correction
291 method proposed by Guidotti et al.¹³ in the context of the very large variations in δ -values
292 encountered when measuring doubly labelled waters (*e.g.* differences of several thousands of
293 ‰ in $\delta^2\text{H}$). Based on the exponential decay of their ME factors with time, which can be fitted
294 with a 3-term exponential function, they concluded that 3 water pools are involved in the ME.
295 The correction equation describes exchange processes between these water pools, constrained

296 by several parameters (e.g. 6 parameters, when assuming 3 water pools), which are adjusted for
 297 each individual run, by minimizing the value for the SD calculated on all injections of the same
 298 sample. This approach better describes the physical processes, and implicitly accounts for the
 299 influences of several successive samples, as long as compositions of the water pools are not
 300 fully renewed. It has been developed in the context of measuring labelled water, and the
 301 parameter adjustment is based on selected samples whose δ -values strongly differ from that of
 302 the preceding samples.

303

304 Here, we propose a simpler correction procedure that is able to account for the long-lasting
 305 memory without the need to adjust any parameters. It assumes the variation with time (*i.e.* from
 306 one injection to the next, since the injections are automated) of 1) the mixing proportions
 307 between the current sample and the residual vapor, following the reference series of ME factors,
 308 and 2) the composition of the residual vapor δ^r . The evolution of the δ^r term was calculated
 309 assuming that, after the last injection of a given sample, its contribution to δ^r follows the ME
 310 factor reference series, whatever the further changes in water samples, until $x_i=0$. Each of the
 311 following samples contribute in turn to δ^r in decreasing proportions as the injections continue.
 312 A mixing equation was thus used to calculate the composition of the residual vapor during the
 313 injection i of the current sample n (δ_i^r) as a function of the compositions of four previous
 314 samples (the number of samples expected to contribute to the residual vapor depends on the
 315 number of injections performed for each sample and of the decay of the ME, and can be
 316 adjusted) :

317

$$318 \quad \delta_i^r = \frac{x_i^{n-1} \delta^{n-1} + x_i^{n-2} \delta^{n-2} + x_i^{n-3} \delta^{n-3} + x_i^{n-4} \delta^{n-4}}{x_i^{n-1} + x_i^{n-2} + x_i^{n-3} + x_i^{n-4}} \quad \text{Eq. (2)}$$

319

320 where δ^{n-1} to δ^{n-4} are the compositions of the four samples that precede the sample n and x_i^{n-1}
 321 to x_i^{n-4} are the ME factor series counted from just after the last injection of the sample $n-1$ to $n-4$,
 322 as described in Figure 3. Practically, this composition is calculated in the data processing
 323 script by using 4 series of ME factors associated to 4 successive samples. The initial δ^r value
 324 is set to the composition of the first sample vial, therefore the number of injections of this
 325 conditioner sample need to be properly chosen.

326

327 Then, the measurement of the current sample n is corrected for the ME using the classical
 328 mixing equation:

329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361

$$\delta_{i,corr}^n = \frac{\delta_{i,meas} - x_i^{n-1} \delta_i^r}{1 - x_i^{n-1}} \quad \text{Eq. (3)}$$

Nevertheless, the application of this equation differs from its classical use, since both the terms δ_i^r (calculated from Eq. 2) and x_i vary from the first to the last injection of a sample. Similarly to the method proposed by Guidotti et al.,¹³ this simplified correction scheme is able to account for several preceding samples. Nevertheless, it doesn't need to make any assumptions on the number of water reservoirs involved in the ME, and it doesn't require to use parameters experimentally determined from the data being considered for ME correction.

3-2-3 Evaluation of the ME correction method

To evaluate the magnitude of the ME and the efficiency of the new correction procedure, we used three runs, which were initially designed to evaluate the short-term performances of the instrument (Table 4). They included blocks of 6 vials filled with the same water, treated as unknown samples, each block separated from the others by a group of 3 WS dedicated to calibration. The sequence was run three times, first with ICE as unknown samples, secondly with TAP and the last run with SEA. Ten injections were performed per vial. The vials placed immediately after the calibration WS group (*i.e.* positions #5, #14 and #23) were the most strongly affected by the compositions of previous samples, while the last vials in each block of similar samples (*i.e.* positions #10, #19 and #28) were not affected by ME.

For all these runs, four different methods were applied to correct the ME, and the calibrated results were compared (Figure 4). In the first method (M1), no ME correction was applied but the first 6 injections of all vials were discarded. In the second method (M2), the data were corrected by considering only the composition of the preceding sample (*i.e.* $\delta_i^r = \delta^{n-1}$). The third data processing (M3) followed the ME correction method proposed by Guidotti et al..¹³ Two water pools were considered, as suggested by the 2-component exponential fit of our ME factors (Figure 1C and 1D). The corresponding 4 parameters were adjusted for each run using the measured data corresponding to the WS used for calibration, which covered the highest difference in δ -values between successive samples, and therefore the highest level of information on the ME. The fourth data processing (M4) was performed with the new correction method described in Figure 3, therefore by calculating δ_i^r with Eq. 2, and correcting injections data using Eq.3. In this case, the first group of WS was discarded from the calibration, and kept

362 to initialize the ME correction. For the last three methods, the first 2 injections after the
363 application of the ME correction were discarded.

364

365 For the first samples in each sub-block (Figure 4A to 4D), a slight offset was found when no
366 numerical correction was applied (M1): on average 0.03 ‰, 0.06‰, 0.79‰ and 0.28‰ for
367 $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ and d-excess, respectively. The simple ME correction (M2) over-corrected the
368 values, leading to a reverse offset of -0.03 ‰, -0.06 ‰, -1.06‰ and -0.57‰ on average for
369 $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ and d-excess, respectively. The third method (M3) led to results almost similar
370 to what was obtained by simply discarding the first injections (M1). The new correction (M4)
371 led to intermediate values, still slightly over-corrected in $\delta^2\text{H}$ but very close to the reference
372 compositions with offsets of 0.00 ‰, -0.01 ‰, -0.40‰ and -0.31‰ for $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ and d-
373 excess. For each method, the highest offsets were found for the lowest δ -value (the ICE WS).
374 The ^{17}O -excess composition was not affected by the ME correction because similar ME
375 coefficients were used for $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ (not shown). However, slight differences were found
376 when discarding the first 6 injections, since a lower number of injections was used for ^{17}O -
377 excess determination.

378

379 The last vials of each block of unknown samples (#10, #19 and #28) were expected to be
380 deprived of the ME, but their calibrated compositions depend on the correction applied to the
381 WS data used for the calibration. When simply discarding the first 6 injections (M1), and with
382 the 2-pool assumption (M3), significant offsets were observed, whose magnitude and direction
383 depend on the isotopic composition of the unknown samples (Figure 4E to 4H). The offset was
384 particularly marked for values close to the VSMOW composition (the SEA WS).

385

386 The poor improvement obtained with the Guidotti et al. approach compared to the method M1
387 came from the criterion used to adjust the correction parameters. The method was developed
388 for measuring spiked water samples which carry a huge ME. In these cases, the ME correction
389 greatly improve the SD calculated for successive injections of a sample. However, our results
390 showed that the raw and ME-corrected groups of injections of a sample did not clearly differ in
391 their SD, although they were different in their absolute values, with regards to
392 the expected accuracy, which is almost one order of magnitude higher than what is expected
393 for labelled water samples.¹³

394

395 3-3 Long-term statistics on routine runs

396

397 The new ME correction was applied retroactively to all sequences from October 2018 to March
398 2020. The data processing followed several successive steps. First, all the injections of a run
399 were ME-corrected. The first two injections of each vial were systematically discarded because
400 of the high variability in the corresponding ME factors (Figure 1G). Then, ME-corrected raw
401 values of $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ were estimated for individual WS vials by averaging the remaining
402 injections. The calibration equations were established using the average of all replicates of the
403 3 WS and applied to all the injections of a run. Finally, the calibrated data were averaged for
404 each vial to obtain $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$, and thereafter ^{17}O -excess and d-excess.

405

406 We observed slight differences between the compositions of the two QA replicates. The
407 composition of the QA sample placed after the first group of WS was systematically lower than
408 that of the second replicate in the run architecture #2. The difference between the 2 replicates
409 (0.03 ‰ for $\delta^{18}\text{O}$ and 0.29‰ for $\delta^2\text{H}$) remained in the range of the final long-term
410 reproducibility, but this imperfect ME correction showed that careful attention should also be
411 paid to the run architecture to minimize the magnitude of the ME.

412

413 Despite very good linear regression coefficients (>0.999995), the linearity of the calibration
414 was not fully reached, since the scaling factors based on the ICE-TAP WS (0.969, 0.978 and
415 0.971 for $\delta^{17}\text{O}$, $\delta^{18}\text{O}$ and $\delta^2\text{H}$), were found slightly different from those based on the TAP-SEA
416 intervals (0.975, 0.984 and 0.978 for $\delta^{17}\text{O}$, $\delta^{18}\text{O}$ and $\delta^2\text{H}$). This slight non-linearity remained
417 lower than what could result from the bias due the ME. In addition, the 3-point calibration was
418 compared to a normalization performed using only the two most contrasting WS (ICE and
419 SEA). The differences in calibrated data remained lower than 0.01‰ for $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ and
420 lower than 0.2‰ for $\delta^2\text{H}$.

421

422 Long-term statistics were calculated for the two run architectures used for routine analysis,
423 using the average composition of QA replicates per run (Table 5). In addition, individual
424 compositions of all QA are plotted for the entire measurement period (Figure 5). The resulting
425 precision (estimated as the SD of QA compositions) was slightly improved with the run
426 architecture #2 (Table 5: 0.015‰, 0.023‰, 0.078‰ and 0.006‰ and 0.173 ‰ for $\delta^{17}\text{O}$, $\delta^{18}\text{O}$,
427 $\delta^2\text{H}$, ^{17}O -excess, and d-excess respectively), and fell within the range of the instrument

428 specifications, corresponding to the continuous measurement of the same water. In addition,
429 the accuracy estimated as the deviation (bias) from the expected composition (Table 5) was
430 also improved for the run architecture #2: 0.005‰, 0.015‰, -0.124‰, -0.002‰ and -0.252‰
431 for $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$, ^{17}O -excess, and d-excess respectively. These offsets were lower than the
432 SD for each tracer, except for d-excess, as a result of opposite, though low, offsets for $\delta^{18}\text{O}$ and
433 $\delta^2\text{H}$.

434

435 **4- Discussion and conclusion**

436

437 Our results demonstrate the significant level and persistence of the ME when analyzing water
438 samples with WS-CRDS. For example, the contribution of residual vapor is still between 0.5%
439 and 1% in the 10th injection of the same water for oxygen and hydrogen. It is thus essential to
440 apply a robust correction during post measurement data processing to obtain high precision
441 isotopic data. A similar impact of the ME was evidenced for $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$, allowing to use the
442 same correction for the two isotopic species. Consequently, the application of a ME correction
443 is transparent for the ^{17}O -excess determination, and affects neither the value nor the
444 reproducibility of ^{17}O -excess data.

445

446 We showed that the ME required up to 45 successive injections of the same sample to be
447 removed, with respect to the instrument precision, and to the δ -value intervals corresponding
448 to our WS. This corresponds to about 5 hours of measurement with the “high precision mode”
449 coordinator and an injection volume of 1.6 μL . This duration is longer than usually expected⁴⁻
450 ^{6,9} and has two main consequences. First, it makes the systematic evaluation of the ME in
451 routine runs questionable, as it would be too time-consuming for a full assessment. With regards
452 to the relative stability of the ME with time, it is preferable to perform a robust determination
453 of ME factors based on several dedicated runs to obtain a robust average series. Second, it
454 implies that several successive samples contribute to the composition of residual vapor
455 affecting the measurement of a given sample in routine runs. Such a long-term persistence of
456 the memory was already pointed out and accounted for in the 3-pool-exchange model proposed
457 by Guidotti et al. (2013) for memory correction of doubly labelled waters characterized by huge
458 differences in δ -values.¹³ We proposed here an alternative ME correction method based on a
459 mixing equation to calculate the composition of the residual water vapor as a function of the
460 compositions of a number of previous samples. Compared to the method proposed by Guidotti

461 et al., this simplified correction scheme was found more suitable for natural samples, for which
462 the magnitude of the ME is lower, and the expected precision higher than for labelled water.

463

464 The insufficient number of injections used to assess the ME in most current practices^{4-7,9} leads
465 to the underestimation of the full magnitude of the memory. Nevertheless, the influence of an
466 inappropriate ME correction may remain hidden in a number of situations, *e.g.* in cases where
467 the under-estimated ME is compensated for by the alternation of variable sample compositions
468 mitigated long-lasting ME, *e.g.* when δ^{n-2} is closer to δ^n than δ^{n-1} . However, in a monotonic
469 succession of sample compositions (either increasing or decreasing δ values) the under-
470 estimated ME results in inaccurate data. The new correction formula led to better precision
471 whatever the distribution of sample compositions in the run architecture. A systematic offset in
472 calibrated data may also be hidden when a constant order is adopted in the routine runs for WS,
473 since the calibrated data may be indirectly affected by a bias in the measurement of the WS
474 used for calibration. This effect, shown in Figure 4E to 4H, may remain negligible for sample
475 compositions falling in between WS values. However, significant bias were found for samples
476 close to WS compositions.

477

478 The determination of robust ME factors depends on the assumption of a memory-free value for
479 the composition of the current vial. However, the asymptotic decay of the ME makes difficult
480 to assert a complete disappearance of memory. We found that 50 injections were a good
481 compromise between memory dampening and the possible instrumental drift, with regards to
482 the measurement precision. Nevertheless, this major pitfall combined with the natural
483 variability of ME factors (Figure 1) implies that a perfect correction can never be accomplished,
484 and that it is necessary to minimize the magnitude of the ME by adopting an appropriate run
485 architecture. The magnitudes of the ME (estimated as non-corrected *minus* corrected calibrated
486 data) were compared for two runs of precipitation samples, conducted with the two
487 architectures described previously. Precipitation samples usually encompass a large range of
488 isotopic compositions; in these cases, values varied from $\delta^{18}\text{O} = -8\text{‰}$ to 0‰ . In the run
489 architecture #1, the magnitude of the ME reached 0.07‰ , 0.15‰ , 1.7‰ and 0.7‰ for
490 $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ and d-excess respectively. In the run architecture #2, the differences were
491 reduced by half, remaining below 0.04‰ , 0.06‰ , 0.9‰ and 0.7‰ for $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$ and d-
492 excess respectively. The dampening of the ME was obtained by performing 15 injections of
493 each WS used for calibration, including conditioning samples before each WS group, and

494 systematically analyzing replicates of WS and samples, the replicates being sorted in reverse
495 order in a run sequence. The symmetrical organization also eliminated the need for a drift
496 correction, as this effect is implicitly corrected by averaging replicates. In recent instruments
497 the drift generally remains low (e.g. Schauer et al.⁸ did not evidence any drift during periods of
498 10 consecutive days), but when a drift correction is required, a robust memory correction is
499 essential to avoid the misattribution of the ME to a drift effect.

500

501 The use of a QA WS included in all runs is essential. To avoid the hidden effect of a constant
502 shift, the QA should be placed randomly among unknown samples. In addition, a Quality
503 Control (QC) water would be required to fully assess the data accuracy (trueness). Then, the
504 composition of the QA/QC water has to be determined independently, using primary
505 international standards, with a run architecture different from that of routine measurements, and
506 with a number of injections sufficient to eliminate the ME. Laboratories using IRMS techniques
507 in parallel with laser technology have the opportunity to perform independent measurements.
508 However, the precision obtained for $\delta^2\text{H}$ using IRMS techniques hardly reaches 0.1‰,^{16,20}
509 which would be necessary with respect to what is expected in laser data. Ultimately, dedicated
510 inter-laboratory comparisons are necessary to fully evaluate the data accuracy.

511

512 The difference between the 3-point calibration used in this study and a normalization performed
513 using the two most contrasting WS (ICE and SEA), evaluated on calibrated data, remained
514 lower than 0.01‰ for $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$ and lower than 0.2‰ for $\delta^2\text{H}$. This would plead for the 2-
515 point calibration classically recommended (REF), with a large range of compositions for the
516 WS, since it has the advantage of minimizing error propagation during calibration. However,
517 the unexpectedly long duration of the ME underlines the need for an appropriate correction to
518 compensate for the persistence of a highly different residual vapor during the measurement of
519 the samples that follow the WS. An alternative would be to choose WS compositions closer to
520 the expected sample compositions (e.g. no strongly depleted compositions, when working on
521 temperate climate regions), and use up to 3 WS for the calibration. But the choice of 3 WS for
522 the calibration is time-consuming. Therefore, the use of 2 WS covering a large range of
523 composition would remain the best choice, provided that the ME is properly corrected.

524

525 Finally, the new ME correction method proposed in this study (Python script in the
526 supplementary information) combined with an appropriate run architecture, allowed us to
527 achieve a long-term precision similar to the instrumental specifications provided by the

528 manufacturer. The standard deviation estimated for the QA after modification of the run
529 architecture (7-month period) showed a precision of 0.015‰, 0.023‰, 0.078‰, 0.006‰ and
530 0.173‰ for $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$, ^{17}O -excess., and d-excess., respectively. The offset from the
531 expected composition of the QA was 0.005‰, 0.015‰, -0.124‰, -0.002‰ and -0.252‰ for
532 $\delta^{17}\text{O}$, $\delta^{18}\text{O}$, $\delta^2\text{H}$, ^{17}O -excess, and d-excess respectively (Table 5). In addition, the precision we
533 obtained for ^{17}O -excess, allowed us to demonstrate the first determination of $\delta^{17}\text{O}$ and ^{17}O -
534 excess compositions for the GRESP certified reference material¹⁸ (Table 2).

535

536

537 **Acknowledgements**

538 This study benefited from the ANR HUMI-17 (ANR-17-CE01-0002), ECCOREV and LABEX
539 OT-Med grants. Many thanks to Clément Piel for fruitful discussions, especially during the
540 instrument set-up.

541

542 **References**

- 543 1. Wassenaar LI, Ahmad M, Aggarwal P, et al. Worldwide proficiency test for routine
544 analysis of d2H and d18O in water by isotope-ratio mass spectrometry and laser absorption
545 spectroscopy. *Rapid Commun Mass Spectrom.* 2012;26(15):1641-1648.
546 doi:10.1002/rcm.6270
- 547 2. Wassenaar LI, Terzer-Wassmuth S, Douence C, Araguas-Araguas L, Aggarwal PK,
548 Coplen TB. Seeking excellence: An evaluation of 235 international laboratories conducting
549 water isotope analyses by isotope-ratio and laser-absorption spectrometry. *Rapid*
550 *Communications in Mass Spectrometry.* 2018;32(5):393-406. doi:10.1002/rcm.8052
- 551 3. Aron PG, Levin NE, Beverly EJ, et al. Triple oxygen isotopes in the water cycle.
552 *Chemical Geology.* Published online December 31, 2020:120026.
553 doi:10.1016/j.chemgeo.2020.120026
- 554 4. Pierchala A, Rozanski K, Dulinski M, Gorczyca Z, Marzec M, Czub R. High-precision
555 measurements of delta H-2, delta O-18 and delta O-17 in water with the aid of cavity ring-
556 down laser spectroscopy. *Isotopes in Environmental and Health Studies.* 2019;55(3):290-307.
557 doi:10.1080/10256016.2019.1609959
- 558 5. Wassenaar LI, Coplen TB, Aggarwal PK. Approaches for Achieving Long-Term
559 Accuracy and Precision of delta O-18 and delta H-2 for Waters Analyzed using Laser
560 Absorption Spectrometers. *Environmental Science & Technology.* 2014;48(2):1123-1131.
561 doi:10.1021/es403354n
- 562 6. Coplen TB, Wassenaar LI. LIMS for Lasers 2015 for achieving long-term accuracy
563 and precision of delta H-2, delta O-17, and delta O-18 of waters using laser absorption
564 spectrometry. *Rapid Communications in Mass Spectrometry.* 2015;29(22):2122-2130.
565 doi:10.1002/rcm.7372
- 566 7. Groening M. Improved water delta H-2 and delta O-18 calibration and calculation of
567 measurement uncertainty using a simple software tool. *Rapid Communications in Mass*
568 *Spectrometry.* 2011;25(19):2711-2720. doi:10.1002/rcm.5074
- 569 8. Schauer AJ, Schoenemann SW, Steig EJ. Routine high-precision analysis of triple

570 water-isotope ratios using cavity ring-down spectroscopy. *Rapid Communications in Mass*
571 *Spectrometry*. 2016;30(18):2059-2069. doi:10.1002/rcm.7682

572 9. van Geldern R, Barth JAC. Optimization of instrument setup and post-run corrections
573 for oxygen and hydrogen stable isotope measurements of water by isotope ratio infrared
574 spectroscopy (IRIS). *Limnology and Oceanography-Methods*. 2012;10:1024-1036.
575 doi:10.4319/lom.2012.10.1024

576 10. Penna D, Stenni B, Sanda M, et al. Technical Note: Evaluation of between-sample
577 memory effects in the analysis of delta H-2 and delta O-18 of water samples measured by
578 laser spectrometers. *Hydrology and Earth System Sciences*. 2012;16(10):3925-3933.
579 doi:10.5194/hess-16-3925-2012

580 11. Holko L. Syringe life and memory effects in isotopic analyses performed by liquid
581 water isotopic analysers - a case study for natural waters from central Europe. *Isotopes in*
582 *Environmental and Health Studies*. 2016;52(4-5):553-559.
583 doi:10.1080/10256016.2015.1090987

584 12. Qu D, Tian L, Zhao H, Yao P, Xu B, Cui J. Demonstration of a memory calibration
585 method in water isotope measurement by laser spectroscopy. *Rapid Communications in Mass*
586 *Spectrometry*. 2020;34(8):UNSP e8689. doi:10.1002/rcm.8689

587 13. Guidotti S, Jansen HG, Aerts-Bijma AT, Verstappen-Dumoulin BM a. A, van Dijk G,
588 Meijer H a. J. Doubly Labelled Water analysis: Preparation, memory correction, calibration
589 and quality assurance for delta H-2 and delta O-18 measurements over four orders of
590 magnitudes. *Rapid Communications in Mass Spectrometry*. 2013;27(9):1055-1066.
591 doi:10.1002/rcm.6540

592 14. Schoenemann SW, Schauer AJ, Steig EJ. Measurement of SLAP2 and GISP 17O and
593 proposed VSMOW-SLAP normalization for 17O and 17Oexcess. *Rapid Commun Mass*
594 *Spectrom*. 2013;27(5):582-590. doi:10.1002/rcm.6486

595 15. Python Software Foundation. Python Language Reference, version 3.7. Available at
596 <http://www.python.org>.

597 16. Olsen J, Seierstad I, Vinther B, Johnsen S, Heinemeier J. Memory effect in deuterium
598 analysis by continuous flow isotope ratio measurement. *International Journal of Mass*
599 *Spectrometry*. 2006;254(1-2):44-52. doi:10.1016/j.ijms.2006.05.004

600 17. IAEA. Reference Sheet for VSMOW2 and SLAP2 international measurement
601 standards. Published online 2009.

602 18. IAEA. Reference sheet for certified reference material - GRESP, Greenland Summit
603 Precipitation water. Published online March 2021.

604 19. Gupta P, Noone D, Galewsky J, Sweeney C, Vaughn BH. Demonstration of high-
605 precision continuous measurements of water vapor isotopologues in laboratory and remote
606 field deployments using wavelength-scanned cavity ring-down spectroscopy (WS-CRDS)
607 technology. *Rapid Communications in Mass Spectrometry*. 2009;23(16):2534-2542.
608 doi:10.1002/rcm.4100

609 20. Donnelly T, Waldron S, Tait A, Dougans J, Bearhop S. Hydrogen isotope analysis of
610 natural abundance and deuterium-enriched waters by reduction over chromium on-line to a
611 dynamic dual inlet isotope-ratio mass spectrometer. *Rapid Communications in Mass*
612 *Spectrometry*. 2001;15(15):1297-1303. doi:10.1002/rcm.361

613

614

615 Table 1: Architectures of the runs used for routine measurements. In the run architecture #1
 616 (used over October 2018 – May 2019) unknown and QA samples are replicated (2 or 3
 617 replicates) in the same order. In the run architecture #2 (used over September 2019 – March
 618 2020) unknown samples are replicated in reverse order.
 619
 620

run architecture #1		purpose		
vial position	identifier			number of injections
1	DI water		conditioner	8
2	ICE		Low δ -value WS for calibration	8
3	TAP		Intermediate δ -value WS for calibration	8
4	SEA		high δ -value WS for calibration	8
5-19	samples		unknown samples, including ROB as QA WS	8
20	ICE		Low δ -value WS for calibration	8
21	TAP		Intermediate δ -value WS for calibration	8
22	SEA		high δ -value WS for calibration	8
23-37	samples		Replicated unknown samples, including ROB as QA WS	8
38	ICE		Low δ -value WS for calibration	8
39	TAP		Intermediate δ -value WS for calibration	8
40	SEA		high δ -value WS for calibration	8
run architecture #2				
vial position	identifier			number of injections
1	ICE		conditioner	8
2	ICE		Low δ -value WS for calibration	15
3	TAP		Intermediate δ -value WS for calibration	15
4	SEA		high δ -value WS for calibration	15
5	ROB		QA WS	8
6-20	samples		unknown samples	8
21	ROB		QA WS	8
22-36	samples		replicated unknown samples, in reverse order	8
37	SEA		conditioner	8
38	SEA		high δ -value WS for calibration	15
39	TAP		Intermediate δ -value WS for calibration	15
40	ICE		Low δ -value WS for calibration	15

621
 622

623
 624
 625
 626
 627
 628
 629
 630
 631
 632
 633
 634
 635
 636
 637
 638
 639
 640

Table 2: Mean compositions of in-house WS normalized to the VSMOW-SLAP scale (all values in ‰). ICE, TAP and SEA were measured in the same runs as VSMOW2, SLAP2 and GRESP, including 50 injections per vial, GRESP being treated as an unknown. ROB was measured from 4 runs including SEA and TAP for calibration. The total number of injections used to calculate the δ -values are 60 for ICE, SEA, TAP and GRESP (4 vials, 15 injections per vial), and 40 for ROB (2 vials with the last 15 injections of 50, and 2 vials with the last 4 injections of 15). The sample size is 4 for all mean values and corresponding SDs. See text for a more detailed description of the runs. Regarding the significantly greater precision obtained for ^{17}O -excess compared to $\delta^{17}\text{O}$ and $\delta^{18}\text{O}$, and following the recommendations of Schoenemann et al.,¹⁴ the $\delta^{17}\text{O}$ value and $\delta^{18}\text{O}$ are reported here with a number of decimals higher than allowed by their absolute precisions, in order to be consistent with the ^{17}O -excess calculation ($^{17}\text{O}\text{-excess} = 1000 \times (\delta'^{17}\text{O} - 0.528 \times \delta'^{18}\text{O})$, with $\delta' = \ln(\delta/1000 + 1)$), when all values are expressed in ‰).

purpose	identifier		$\delta^{17}\text{O}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	^{17}O -excess	d-excess
calibration WS	ICE	mean	-14.1836	-26.7411	-203.05	0.026	10.88
		SD	0.026	0.041	0.53	0.007	0.28
calibration WS	TAP	mean	-4.5285	-8.5926	-59.06	0.018	9.68
		SD	0.017	0.021	0.49	0.006	0.41
calibration WS	SEA	mean	-0.1395	-0.2457	0.32	-0.010	2.28
		SD	0.006	0.018	0.09	0.004	0.19
IAEA reference material	GRESP	<i>assigned composition</i>		-33.40	-258.0		9.32
		<i>cum. Uncertainties</i>		0.04	0.4		0.72
		mean	-17.7784	-33.4487	-257.87	0.025	9.72
		SD	0.020	0.034	0.52	0.005	0.30
QA WS	ROB	mean	-4.0588	-7.7084	-52.75	0.018	8.92
		SD	0.025	0.035	0.04	0.010	0.29

641
 642

643 Table 3: Architecture of the run dedicated to estimating the ME factors and corresponding raw
 644 (uncalibrated) delta values obtained from the average of the last 5 out of the 50 injections of
 645 each vial (all values in ‰).
 646
 647

vial position	identifier	$\delta^{17}\text{O}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	^{17}O -excess
1	ICE	-14.98	-28.18	-205.61	0.004
2	ICE	-15.02	-28.25	-205.92	-0.001
3	SEA	-0.55	-1.20	3.02	0.085
4	SEA	-0.55	-1.18	3.41	0.069
5	TAP	-5.01	-9.64	-57.44	0.088
6	TAP	-5.03	-9.65	-57.39	0.079
7	SEA	-0.56	-1.18	3.50	0.066
8	ICE	-14.99	-28.18	-205.44	-0.007
9	TAP	-5.04	-9.66	-57.69	0.074
10	ICE	-15.02	-28.24	-205.68	-0.008
Corresponding SD					
	ICE	0.02	0.03	0.20	0.006
	TAP	0.01	0.01	0.16	0.007
	SEA	0.00	0.01	0.26	0.010

648
 649
 650
 651

652 Table 4: Architecture of the run used to compare ME corrections (10 injections per vial). The
653 sequence was run 3 times, each of them with one of the WS used to fill the 18 sample vials and
654 treated as unknown samples.
655
656

vial position	identifier 1	identifier 2
1	water	conditioner
2	ICE	WS
3	TAP	WS
4	SEA	WS
5		Sample
6		Sample
7		Sample
8		Sample
9		Sample
10		Sample
11	ICE	WS
12	TAP	WS
13	SEA	WS
14		Sample
15		Sample
16		Sample
17		Sample
18		Sample
19		Sample
20	ICE	WS
21	TAP	WS
22	SEA	WS
23		Sample
24		Sample
25		Sample
26		Sample
27		Sample
28		Sample
29	ICE	WS
30	TAP	WS
31	SEA	WS

657
658

659 Table 5: Long term statistics on the QA WS (ROB) calculated separately for the two run
660 architectures described in Table 1. Bias is calculated as measured *minus* accepted values (all
661 values in ‰).
662
663

	run architecture #1	run architecture #2
n	46	20
$\delta^{17}\text{O}$ bias	0.008	0.005
$\delta^{17}\text{O}$ SD	0.018	0.015
$\delta^{18}\text{O}$ bias	0.021	0.015
$\delta^{18}\text{O}$ SD	0.025	0.023
$\delta^2\text{H}$ bias	-0.214	-0.124
$\delta^2\text{H}$ SD	0.146	0.078
^{17}O -exc bias	-0.002	-0.002
^{17}O -exc SD	0.009	0.006
d-exc bias	-0.381	-0.252
d-exc SD	0.147	0.173

664
665

666 **Figure captions**

667

668

669 Figure 1: Average ME factors (solid lines) with their associated standard deviations (dotted
670 lines: $\pm 1SD$, $n = 24$) calculated for $\delta^{17}O$ and $\delta^{18}O$ (1A), and δ^2H (1B). Same average ME
671 factors plotted on a semi-logarithmic scale, with a 2-component exponential fit (1C, 1D).
672 Example of raw data obtained for the low δ -value WS (ICE) following 50 injections of the high
673 δ -value WS (SEA) (1E, 1F). The continuous lines represent the upper and lower limits of their
674 uncalibrated composition, assumed deprived of ME (average of the last 5 injections),
675 considering the instrument precision provided by the manufacturer ($\pm 0.025\%$ for $\delta^{17}O$ and
676 $\delta^{18}O$, and $\pm 0.1\%$ for δ^2H). Comparison of the standard deviations of each series of average ME
677 factors (1G).

678

679 Figure 2: ME factors corresponding to the injections #3 to #8. Average values and
680 corresponding SD (error bars) are calculated for each of the 4 runs performed between
681 September 2019 and February 2020, while the first run performed in 2018 included only one
682 series.

683

684 Figure 3: Schematic description of the build-up of residual vapor (δ_i^r), following Eq. 2, during
685 the injections of a sample of composition δ^n

686

687 Figure 4: Comparison of offsets remaining after different ME corrections (Methods M1 to M4;
688 see text for details). Offsets were estimated from the 3 runs described in Table 4, as the
689 deviation from expected values calculated for vials #5, #14 and #23 (4A to 4D) and for vials
690 #10, #19 and #28 (4E to 4H)

691

692 Figure 5: Calibrated composition of all replicates of QA samples (ROB) included in routine
693 runs between October 2018 and March 2020. Also shown are the reference compositions (black
694 lines) and the average composition of the QA d samples corresponding to the two run
695 architectures (red dashed lines)

696

697

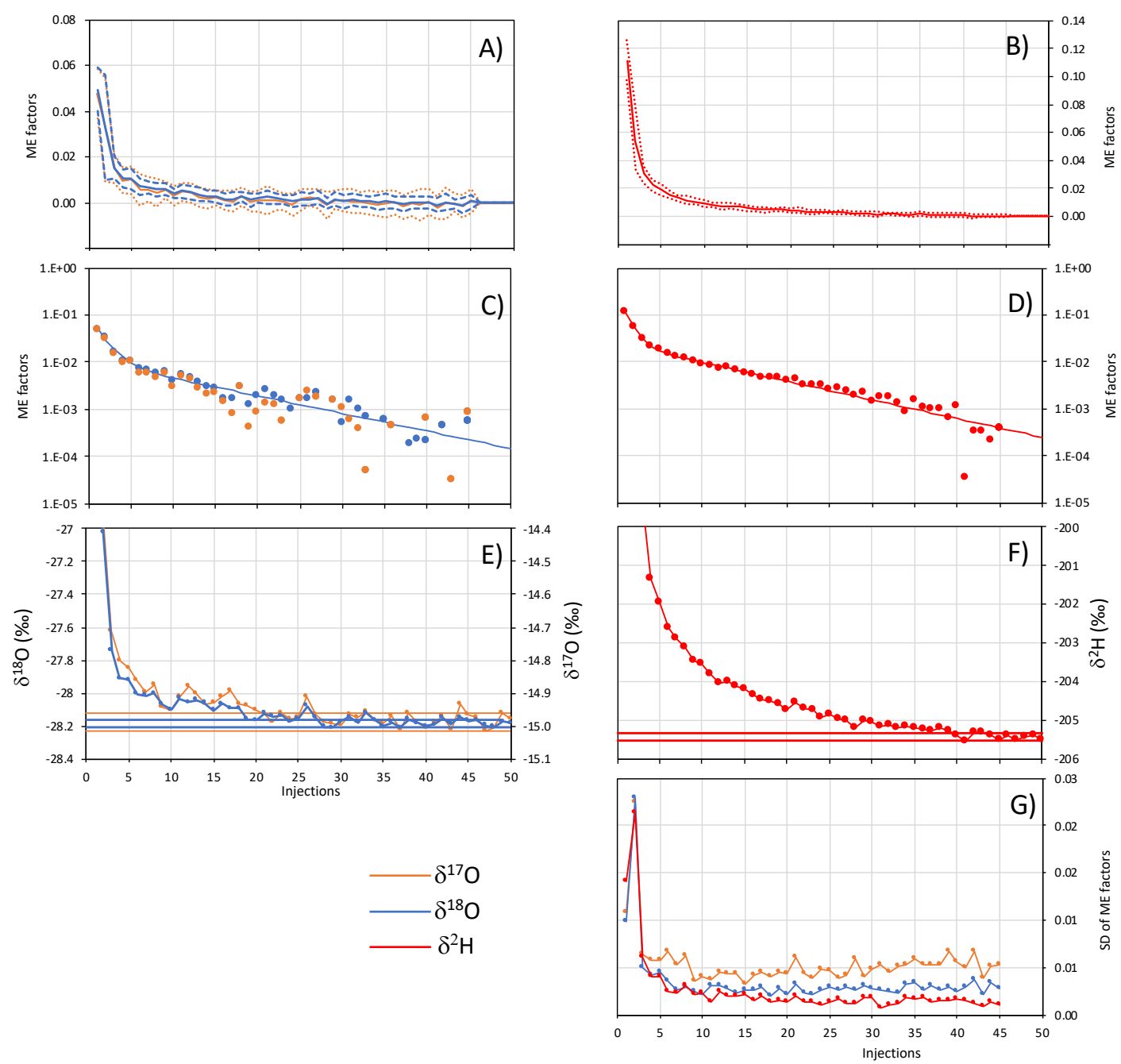


Figure 1

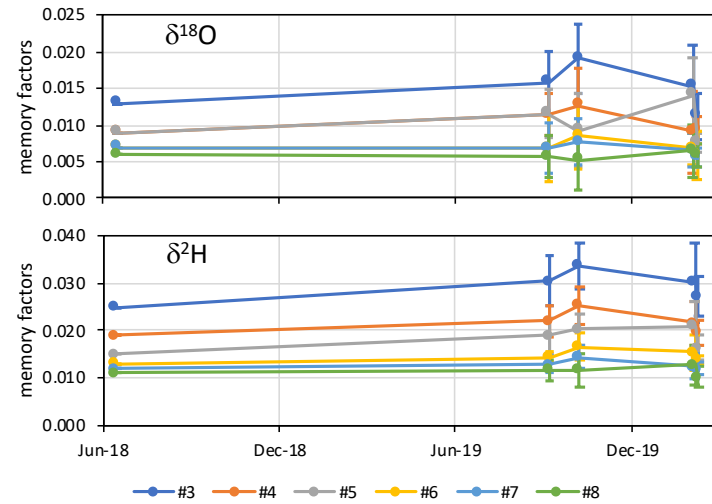


Figure 2

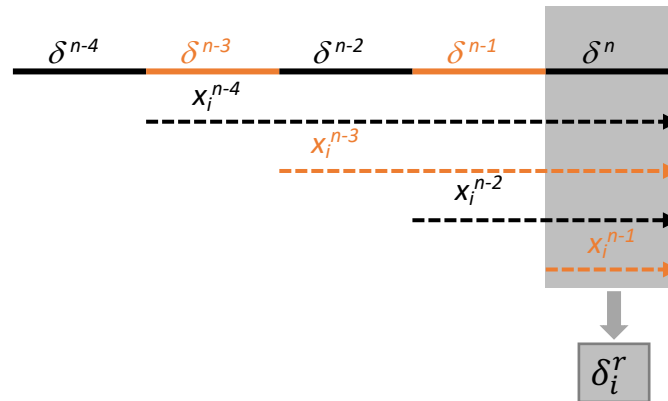


Figure 3

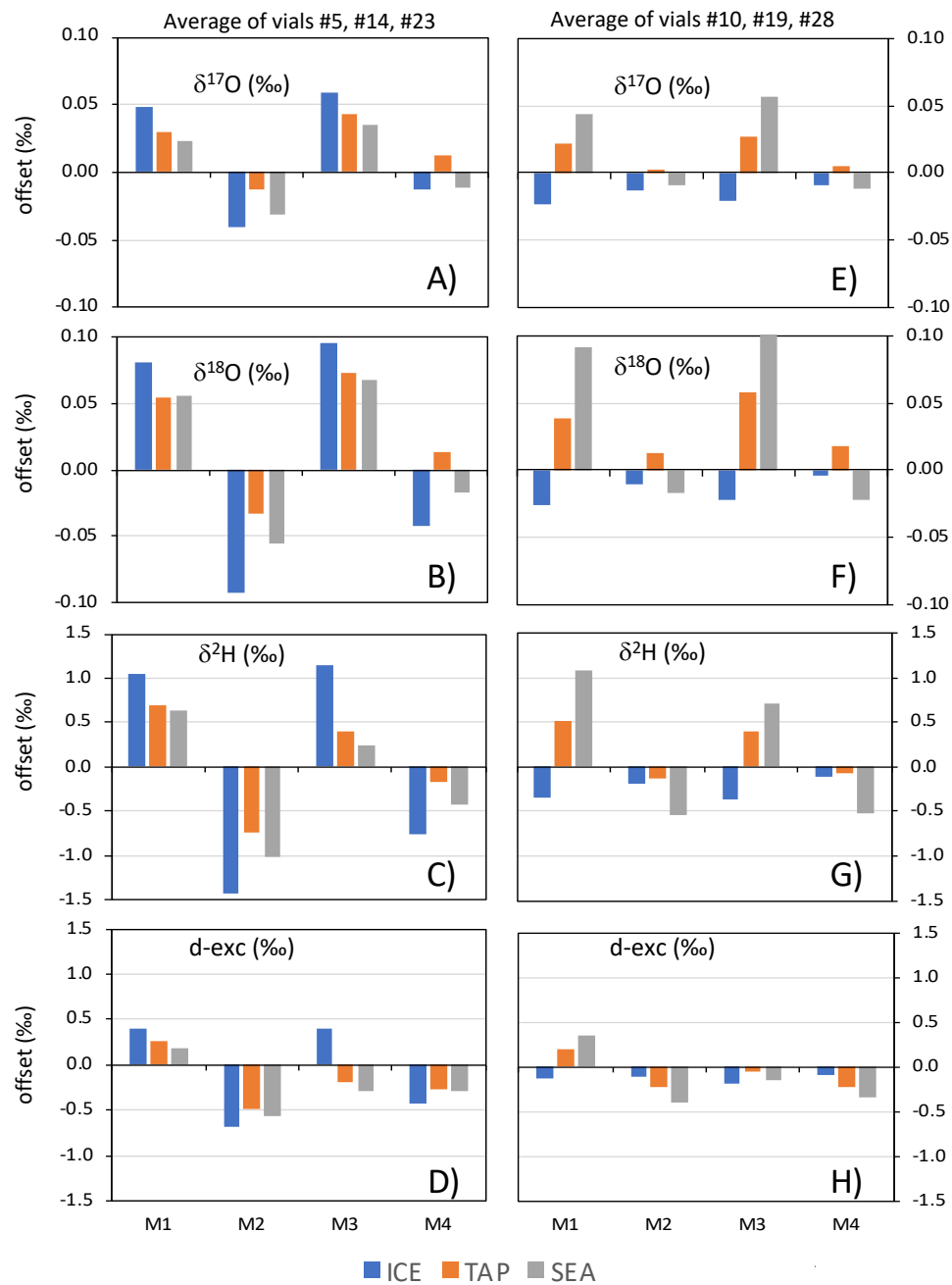


Figure 4

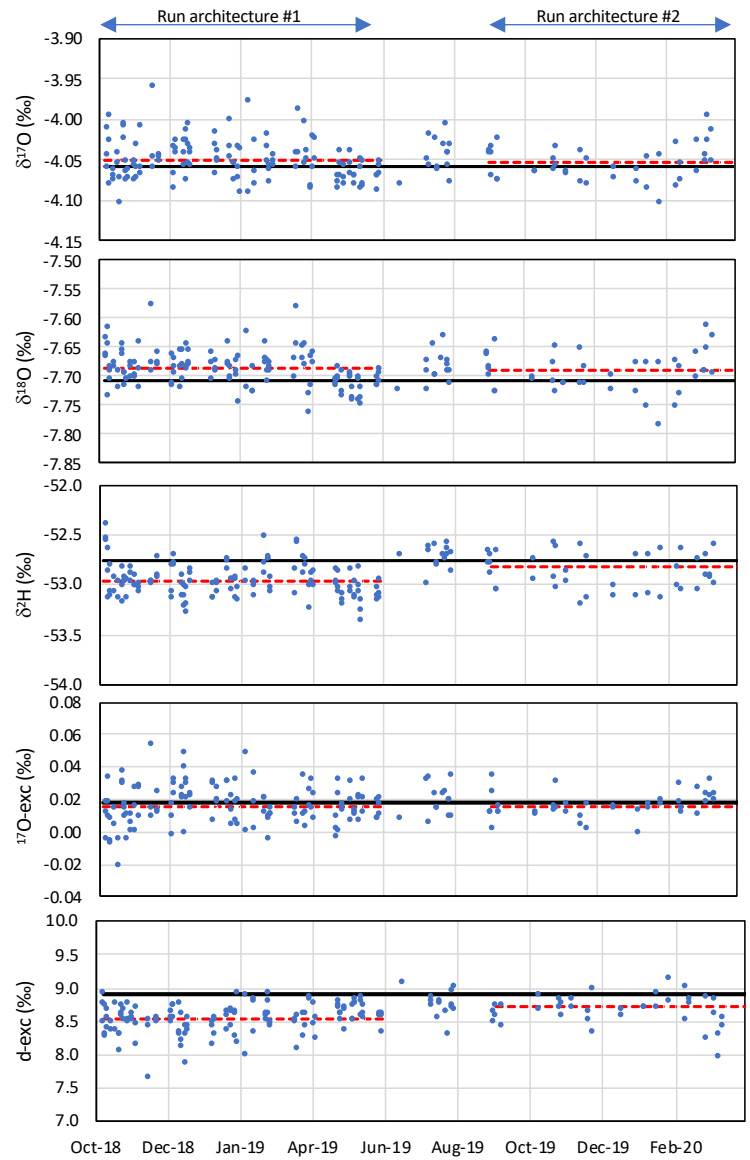


Figure 5