TEL Technical Note No. 02

Performance assessment of stable isotope ratio measurements for highly enriched and highly depleted water samples by laser spectrometry

Manfred GRÖNING

Terrestrial Environment Laboratory, International Atomic Energy Agency, Vienna, Austria

m.groening@iaea.org

May 2018 (updated from Dec 2010)

Abstract

Extensive tests of a Picarro laser system for water measurements are reported with special focus of samples as measured for quality control purposes. A comparison with mass spectrometry methods and other laser spectroscopy instruments was performed, particularly dealing with accurate calibration of laboratory standards, providing an important benchmark performance test.

A special focus was given on the possible extended stable isotopic range for further applications: δ^2H measurements ranged from -990‰ up to +26000‰, those for δ^{18}O from -900‰ up to +50‰. Efficient means were developed for dealing with memory at laser systems for such large isotopic differences, and allowed to evaluate, correct and use all injections starting from the first one, therefore optimising the use of machine time and the achievable precision and accuracy for a given measurement set.

1. Introduction

This Technical Note describes details of the performance of a Picarro laser system for high precision water analysis. A water stable isotope measurement system of type Picarro L1102-i water isotopic analyser using infrared laser light and cavity ringdown spectroscopy was initially tested in the IAEA Isotope Hydrology Laboratory (IHL) during March-April 2009 and from October 2009 onwards. The main purpose was to evaluate the performance of the system for water stable isotope measurements under tight quality control requirements. Tests of laser systems for other applications are reported elsewhere [1],[2],[3]. The IHL laboratory at that time used several different instruments for stable isotope measurements in the area of water resource management, including dual-inlet mass spectrometry using water-gas equilibration, continuous-flow mass spectrometry using high temperature pyrolysis and a further infrared laser spectroscopy method using off-axis integrated cavity output spectroscopy. Thus a method comparison was possible under such a heterogeneous laboratory setup.

The special focus of this Technical Note was given on the possible extended isotopic range for applications and the evaluation of possible precision and accuracy: The range of δ^2H measurements covered values from -990‰ up to +26000‰, those for δ^{18}O ranged from -900‰ to +50‰. Efficient means were required to deal with memory effects at such large isotopic differences. Those means were developed and were applied to routine measurements, allowing to evaluate, correct and use all sample injections without the need to discard leading injections. This was optimising the use of machine time, provided an increased understanding in efficient memory correction and improved the achievable precision and accuracy.

In this respect especially interesting is the measurement precision showing the degree of data scatter around a mean value, and the accuracy showing the residual bias from previously known reference values.

An important topic with the advent of new analytical techniques is the challenge for many laboratories for analysis of stable isotopes with different instrumentation and variable precision levels and the combination of such heterogeneous data sets. It is particularly important to choose best strategies to combine all such data for an analyzed sample to derive one best value to be reported by the laboratory. It is evident that simple approaches like the arithmetic mean will not work under such circumstances due to variations in precision and number of analyses. The final aim is to achieve comparability, so that isotopic values derived from any instrumentation are in accordance with results from other methods within their respective uncertainties.

2. Original instrument set up and changes applied later on

The Picarro L1102-I analyzer was equipped with a vaporizer unit for liquid water injection, and with a HTC-Pal autosampler, CTC Analytics, with two sample trays of 54 positions each for automated injection.

The instrument was tested with various syringes (volumes of 5 µL and 10 µL gas-tight) purchased from two different companies (SGE Analytical Science and Hamilton). It was decided to use 10 µL gas-tight syringes type 10F-C/T-GT-5/0.63, SGE (P/N 002987) in routine mode, as they seemed to be most reliable in operation (several other syringes failed in brand-new stage as the piston did not move smoothly). As a syringe holder for 10 µL syringes was not recognized by the used autosampler, a 5 µL holder was used and the programmed injection amount adjusted. Syringes were initialized before use by flushing with distilled water. During routine operation the syringe was first injecting twice 2 µL sample water into the waste port, then performing three pre-flushes (piston moving three times up to 5 µL level), before the actual injection of 2 µL into the vaporizer was performed.

Septa used for the water injections were of type Auto-Sep 9.5mm, SGE Analytical Science (P/N 041871), and easily could operate more than 1000 injections each (and were changed earlier in routine mode after latest 800 injections).

As supply of dry carrier gas, in first instance a high pressure nitrogen gas cylinder was used with appropriate pressure reduction to 1200 hPa. Later on it was replaced by a small membrane pump, sucking in laboratory air, with attached pressure fine tune valve and a drierite column to ensure a constant pressure dry gas supply at 1200 hPa. The residual water vapour pressure signal for the dry carrier gas in the Picarro analyzer was about 50 ppm for the high-pressure gas cylinder and about 120 ppm using the membrane pump and drierite column. Each drierite column could be operated for more than two weeks before it had to be reconditioned in an oven at 250 °C for three hours.

The volume for each water injection was set to 2 µl water, resulting in a corresponding response signal of about 22000 ppm in the analyzer.

The number of individual injections for each sample was normally set to 6, and to 18 in some cases for extensive tests of memory effects. During tests all samples were analyzed several times per run to generate as much information on system performance as possible. In three special tests highly enriched samples (δ^2H varying from +1000 to +26000 ‰) were used followed by more than 100 injections of a well known standard, in order to get a best mode evaluation of memory effects and allow for efficient memory correction in routine samples.

At the beginning of any analytical run, a “dummy” sample is measured to initialize the measurement system. The results of this sample are otherwise ignored, except allowing for memory correction of the first real sample.

3. Overview on performed tests

An extensive series of tests was performed using the available Picarro L1102-I analyzer. The whole isotopic range covered in these experiments is shown in Fig.1, ranging from samples nearly free of heavy isotopes (-998‰ for δ^2H and -900‰ for δ^{18}O) to samples heavily enriched in δ^2H (up to 26000‰).

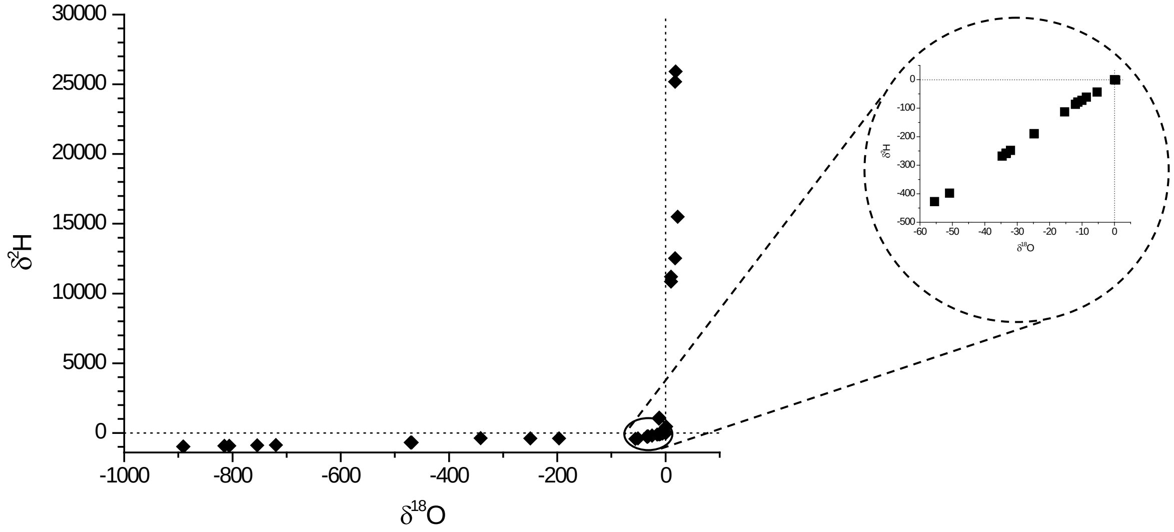

Figure 1: Range of isotopic compositions of samples covered in these experiments shown in an extended δ^2H/δ^{18}O plot. The typical range of an ordinary δ^2H/δ^{18}O plot including VSMOW2 and SLAP2 as end-members is shown in the enlarged inset on the right side.

Two basic memory tests were performed using on one hand isotopically enriched samples (up to 26000 ‰ for δ^2H) and on the other hand isotopically depleted samples (about -998 ‰ for δ^2H and -900 ‰ for δ^{18}O), see Figure 1. The main reason for these tests was to understand and to determine precisely the occurring memory effects, in order to be able to efficiently correct routine samples. Additionally those tests demonstrated the analytical suitability of the instrument over a wide range of isotopic compositions.

Further experiments performed ranged from the reproducibility test for analyses of various samples during a day to long term reproducibility of a sample over more than one year. They included scheduled calibrations of internal laboratory standards versus VSMOW2 and SLAP2. Other tests included the calibration of a series of over 50 external laboratory standards by use of different instrumentation. The latter two tests included internal comparability tests among available stable isotope measurement equipment. In the following sections, some results of the individual tests are described.

4. Measurement evaluation procedure:

4.1 Used evaluation settings and tools:

For a routine measurement of a water sample, six injections are performed for each sample. This is a compromise necessary due to the memory affected first injections and a time consuming effort in more measurements. In several cases, 18 consecutive injections were performed to better assess and correct the memory effect of the instrument. The first sample in any measurement run was considered as “dummy” sample, just to prime the instrument, and was not evaluated further.

The data evaluation was performed by SICalib, an Excel spreadsheet using VBA macros to perform a set of evaluation steps: memory correction, drift correction, calibration of measured data with two laboratory standards, and evaluation of the combined uncertainty for each measurement [4]. In this process no individual data are rejected per default (e.g. the first injections due to memory effects), but rather corrected for the effects, and all data used finally. After initial automated evaluation, various manual refinements can be performed to check and eventually further improve the data consistency.

Finally at the IHL laboratory, all individual calibrated data and uncertainties of all injection were then imported into the SQL database LabData [5], which was used at that time for management of all analytical data of the laboratory. In the database a weighted mean was calculated from all individual measurements of any sample using their individual combined standard uncertainties, and therefore all available data from all instrument used for water stable isotope measurements could be combined, regardless whether these were analyses from mass spectrometers or laser instruments.

There are few mandatory requirements for the setup of any analysis run. Two laboratory standards are mandatory to be used for proper data calibration and normalisation. A third sample is used as quality control sample and should be included. At least the two standards used for calibration have to be measured at least twice during a daily run to allow for an assessment of possible drift-effects. In general, three to four individual samples of each standard should be included in each run, resulting in a number of measurements which reduces the overall uncertainty of those standard measurements. There are no special requirements on predefined sample standard sequences or individual positions for any sample in a given analysis run.

In the evaluation, five uncertainty components are combined according to the Guide of the Expression of Uncertainty in measurements (GUM): uncertainties of the calibration values of the laboratory standards are combined with their measurement uncertainties at the given day, and of course the measurement uncertainty of the given sample is evaluated and added. That last component is the most uncertain one, and since its magnitude is of course unknown, care has to be taken not to over- or underestimate its magnitude. Details for the correction procedure are discussed elsewhere [6].

The evaluated combined uncertainty for each injection is robust, since it is based on the evaluation of all measurement data and is not potentially biased by an unreflected discarding of first injections. Due to the use of all measurements in the uncertainty evaluation procedure, the risk of an unrecognised bias is minimised, as by choosing wrong correction parameters, the corrected results would immediately reflect an increase in their standard deviation versus uncorrected data, which of course should not be the case.

4.2 Memory correction:

The most important individual data evaluation procedure is dealing with sample-to-sample memory. This memory contribution is quite significant for the Picarro instrument, and has to be taken care of and corrected for. There are at least three basic procedures to handle the memory effect, which are discussed below:

- Deletion of first few injections for each sample:
  The easiest solution is just to ignore the first two or three (Picarro default) measurements for each sample and to take the remaining ones as being unaffected from memory. While this is the easiest possible action, its underlying assumption is not true that the remaining injections would be not affected by sample-to-sample memory.

- Assessing the memory effect from daily measurements itself and from stored previous data:
  The approach taken in SICalib requires (infrequent) analysis of 18 injections for two samples with distinct isotopic difference. From those data the memory contribution for each of the 18 injections can be modelled and calculated. Those memory correction factors are stored in a database which can be used for any routine run, or they can be measured and calculated separately within each routine run. This allows to minimise the additional measurement time needed (4 hours in total for the additional 12 injections for two samples, if added to each run). Each injection result can then be memory corrected by use of the corresponding memory factor and the isotopic difference to the previous sample. This procedure significantly outperforms method a) and has the additional benefit that no injection is being discarded and therefore the statistical significance of produced data is improved.

- Assessing the memory effect from isotopically spiked samples:
  This method is used here to provide unbiased memory correction factors. Method b) needs to assume that the last measurements are not memory affected, since only then the isotopic difference to the previous sample can be calculated. This is true for normal samples with small isotopic differences, but not for larger isotopic differences as often encountered for different laboratory standards embracing the whole range of typical samples in a laboratory. In this version, an isotopically enriched sample is measured for at least 36 injections, and then a well known laboratory standard is measured for an extended number of injections, up to 180 injections in the case of the highly deuteriated sample. Then the real memory correction factors for the significant first 6 or 18 injections can be calculated and be applied for all routine measurements (see the discussion in the next section).

The basic requirement is to plan each measurement sequence to such detail, that the most appropriate memory correction can be applied (e.g. define the needed number of injections according to expected isotopic difference between samples). There should be only rare cases when more than 18 injections would be needed in case of isotopically marked samples.

5. Performance tests:

5.1 Experiments with δ^2H enriched water:

Few years ago an enriched water standard was prepared in the laboratory following a Safeguards request as a dilution of a commercial deuteriated water standard (Sigma-Aldrich, 99.96% 2H) with an internal laboratory standard Std12 (δ^2H = -86.4 ± 0.3 ‰; δ^{18}O = -12.03 ± 0.02 ‰). Measurements by dual inlet mass spectrometry resulted in a δ^2H value of 1033.4 ± 0.5 ‰ (n=6). A gravimetric mixing calculation resulted in a nominal value of 1033.9 ‰ in excellent agreement with the measurement and confirming the purity of the commercial deuteriated standard. The (unknown) δ^{18}O value of the enriched standard has a minimal influence on the δ^2H value of the mixture (a positive 1000 ‰ shift in δ^{18}O would cause only a marginal shift of 0.3‰ for δ^2H of the mixture, and such a large δ^{18}O shift can be excluded from δ^{18}O measurements of the mixture).

Measurements performed by the Picarro instrument indicate a δ^2H value of 1033.0 ± 0.3 ‰ (n=18) in good agreement with the previous results.

A series of other enriched samples (partly samples from electrolytic tritium enrichment) were analyzed with δ^2H values of up to 26000 ‰. Measurements of such highly enriched samples cause a longer persisting contamination of the instrument by memory effects, which can be efficiently quantified by this method.

18 injections of a highly enriched sample (26000 ‰ in δ^2H) were used to determine the memory effect on a subsequently analysed well known sample (Std8), which was measured afterwards with 180 consecutive injections (Figure 2). After 18 injections of the second sample, the remaining memory effect had decreased to 0.25% (of the isotopic difference between the samples), still corresponding to a 70‰ offset in δ^2H. After 180 injections the memory had decreased to 0.024%, still causing an offset above 6‰. It is certainly not advisable to perform such measurements in routine mode due to the prevailing contamination, but they provide valuable information on the long term tailing and they can be used to determine appropriate memory correction parameters, which allow to develop very efficient memory corrections for routine samples. That allows in routine mode to correct all injections reliably, and then allows to use all corrected measurements for data evaluation (no discarding of any injection result necessary anymore). This therefore optimizes the use of measurement resources, improves statistical approaches due to higher number of analyses, and finally improve achievable precision and accuracy for a given number of measurements.

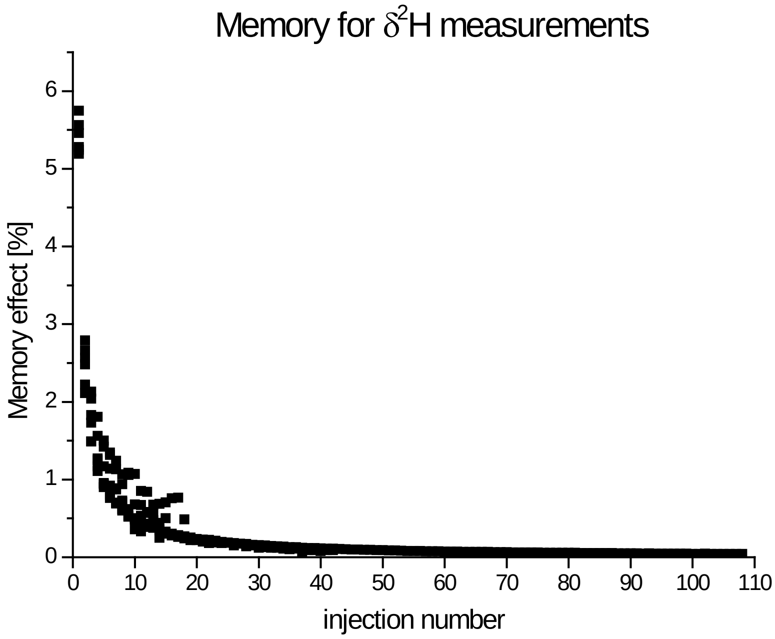

Figure 2: Relative memory effect as percentage contribution of δ^2H-value differences of two samples (denoted as $\Delta\delta^2$H) to injections of the second sample. Displayed are significant memory contributions for the second measured sample for five data sets covering $\Delta\delta^2$H ranges between samples of -26000 ‰ to +1100 ‰. For details see text.

A similar procedure can be used for δ^{18}O, in that case with higher data scatter due to the much smaller isotopic differences, approaching very fast an insignificant level indicated by high data scatter and approaching the measurement uncertainty (Figure 3).

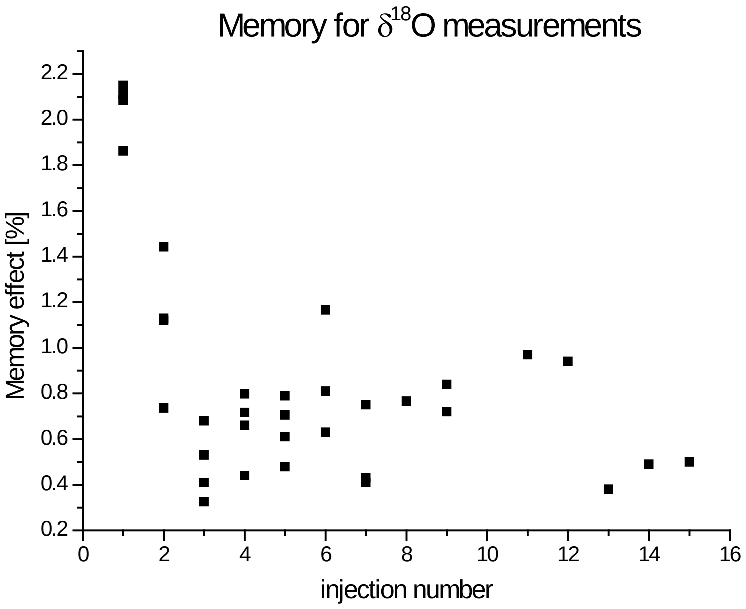

Figure 3: Relative memory effect as percentage contribution of δ^{18}O-value differences of two samples (denoted as $\Delta\delta^{18}$O) to injections of the second sample. Displayed are significant memory contributions for the second measured sample for five data sets covering $\Delta\delta^{18}$O ranges of -26 ‰ to +42 ‰. For details see text.

Determination of the memory correction parameters enables the calculation of optimal number of injections for a given set of samples and standards. For example, if a given set of laboratory standards has an $\Delta\delta^2$H isotopic difference of +200 ‰ (e.g sample isotopic δ-values of 0 ‰ and -200 ‰), then it can be easily calculated with the data from Figure 2 that after 18 injections (still 0.25% memory contribution) the resulting bias is of a magnitude of about 0.5‰, so is at the same level as the 1σ reproducibility for this Picarro analyzer. Thus it will not cause any very notable bias for sample results. Analyzing samples with smaller isotopic spread and using internal laboratory standards closely matching those samples, the number of necessary injections can be further reduced.

It is possible now to elaborate on the necessary number of injection to obtain useful results. After six injections, the relative memory contribution for the sixth injection dropped to about 1%. Therefore without memory correction it would be necessary to limit the δ^2H isotopic difference between consecutive samples to about 50 ‰ to have at least the sixth injection with a memory caused offset of only 0.5 ‰ (and larger offsets for earlier injections). As the isotopic difference between calibration standards is already in the order of 400 ‰, it should be clear that a suitable memory correction is needed to ensure reliable accurate data. Just ignoring the first three injections as proposed by some instrument manufacturers is not sufficient, as injections 4-6 are still considerably affected by the memory and would cause a respective offset/bias versus the true isotopic composition.

It should be noted that the memory correction factors in Figure 2 are normalized and displayed as relative percentage values of the isotopic sample-to-sample differences, and were evaluated by quite different sample combinations, including changes to more positive or more negative δ-values of quite different magnitudes. Despite of that fact, the relative memory correction factors displayed for all such experiments are very consistent.

The scatter among memory correction factors for δ^{18}O measurements (Figure 3) are much larger, which is due to the much smaller range covered (70 ‰ only). Samples depleted in δ^{18}O were not used for this assessment due to only very small amounts available. Still, the averages for the correction factor for the first few injections allow a reasonable memory correction (which gets insignificant later on).

The use of these memory correction factors allows for a very efficient memory correction of the first measurements, as needed for routine samples and the used 6 or 18 injections. At the same time this procedure constitutes no risk for data integrity, as any wrong or overcorrection would result in bad matching of the different injections, which would be automatically reflected in an increase of the standard deviation for all injections of the sample, and in an increase of the combined standard uncertainty of each injection. The user is notified of such a case due to a pre-set warning level for the maximal tolerable combined standard uncertainty and could investigate the case in detail.

5.2 Experiments with δ^2H and δ^{18}O depleted water:

Two extremely deuterium-depleted water samples “M” and “UM” (15ml each) were made available to the IAEA by Neil Sturcio, University of Illinois at Chicago, extending the available δ-range for δ^2H to nearly -1000 ‰ and that for δ^{18}O to -900 ‰.

Small aliquots of the IHL internal laboratory standard “std10” (Southpole water) were accurately weighed into three small glass bottles under dry argon atmosphere. Small aliquots of sample M were taken out of the bottle under dry argon atmosphere and transferred into the glass bottles, to obtain three dilutions of sample M with about 10%, 20% and 50% of std10 admixed (Table 1).

Table 1: All weighings for sample mixtures were performed with a precision balance Kern ABT 120-5DM (+-0.01mg resolution up to weight of 42g):

	Weight sample M [g]	Weight std10 [g]	Fraction of std10 in mixture
M10 (90% M + 10% std10)	4.45692	0.49244	0.099496
M20 (80% M + 20% std10)	3.95121	0.99746	0.20156
M50 (50% M + 50% std10)	2.47254	2.50388	0.50315

Small glass vials were filled with approximately 1 mL of the five samples (M, UM, M10, M20, M50) and of the laboratory standards (std9, std10, std11) under dry argon atmosphere.

Delta-values for the depleted samples were measured directly (values for “M” and “UM” in tables 2 and 3). They were also back-calculated using the delta-values of the three mixture samples (“M10”, “M20”, “M50”) and using the reference value of the standard “std10”. The spread of directly measured values agrees with the spread of back calculated values (Figure 4). Those samples were measured both by the Picarro L1102-i analyzer as well as by three Los Gatos Research DT-100 Liquid Water Isotope Analyzers (LGR). Due to the very low abundance of ^{2}H and ^{18}O in the samples, no attempt was even made to use mass spectrometry for analysis due to the resulting insignificant voltage signals at the minor Faraday cups.

Table 2: δ^2H measurements of depleted samples M, UM and mixtures with std10 in various laser instruments.

Instrument	Std11	Std9	Std10	M50	M20	M10	M	UM
Reference	0.0 ±0.4	-189.2 ±0.2	-397.6 ±0.3					
Picarro	0.0 ±0.4	-189.7 ±2.5	-397.6 ±0.5	-694.45 ±0.5	-866.5 ±1.0	-932.1 ±0.4	-986.4 ±0.6	-988.7 ±0.6
LGR0	0.0 ±0.9	-189.3 ±1.5	-397.6 ±1.9	-697.9 ±0.9	-876.4 ±1.2	-939.9 ±0.6	-998.7 ±0.6	-1000.3 ±1.0

LGR1	0.0 ±2.0	-189.7 ±1.8	-397.6 ±1.4	-695.8 ±0.5	-874.7 ±0.5	-937.7 ±0.4	-996.7 ±0.4	-997.2 ±0.5
LGR2	0.0 ±1.7	-189.2 ±2.1	-397.6 ±1.6	-695.8 ±1.0	-877.0 ±0.8	-938.0 ±0.8	-998.2 ±0.6	-997.6 ±0.4

Table 3: $\delta^{18}O$ measurements of depleted samples M, UM and mixtures with std10 in various laser instruments.

Instrument	Std11	Std9	Std10	M50	M20	M10	M	UM
Reference	0.07 ±0.03	-24.75 ±0.02	-50.89 ±0.03					
Picarro	0.07 ±0.12	-25.00 ±0.62	-50.89 ±0.07	-478.22 ±0.17	-731.97 ±0.49	-823.69 ±0.32	-904.59 ±0.60	-907.51 ±0.33
LGR0	0.07 ±0.30	-24.76 ±0.41	-50.89 ±0.56	-474.35 ±0.42	-730.15 ±0.54	-817.88 ±0.21	-902.78 ±0.25	-903.65 ±0.39
LGR1	0.07 ±0.24	-24.73 ±0.29	-50.89 ±0.56	-475.57 ±0.22	-733.29 ±0.47	-821.89 ±0.13	-907.88 ±0.19	-908.40 ±0.33
LGR2	0.07 ±0.29	-24.69 ±0.40	-50.89 ±0.77	-475.18 ±0.35	-732.40 ±0.62	-821.28 ±0.20	-907.00 ±0.30	-907.57 ±0.44

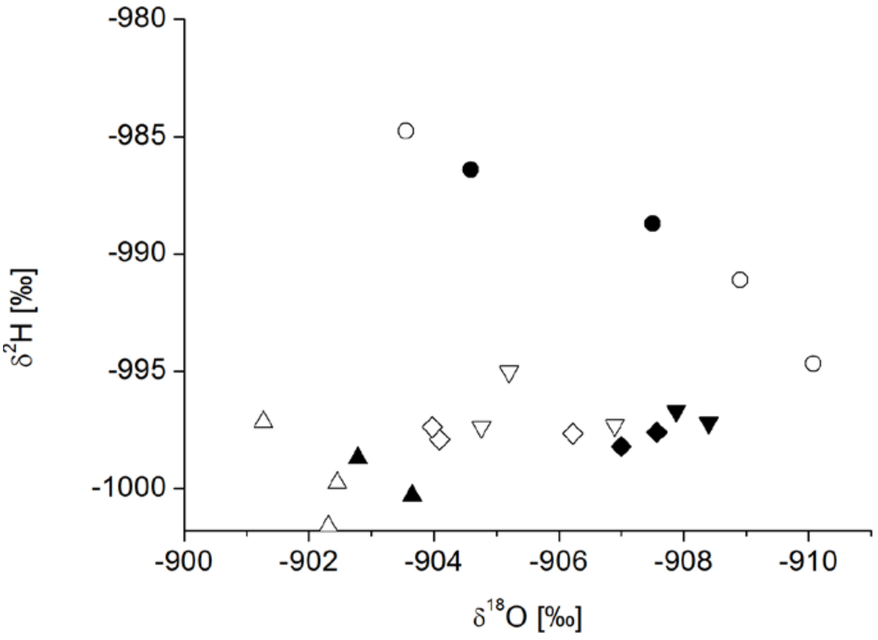

Figure 4: Delta-Plot of water samples very depleted in δ^2H and $\delta^{18}O$, measured by four different laser systems (Picarro as circles, LGR as triangles and squares), measured directly (filled symbols) and back-calculated from dilutions with 10%, 20% and 50% of isotopically well known laboratory standard (open symbols). Data from Picarro were subject to analytical problems due to leaky valve.

The data in Figure 4 show a good correspondence of direct δ^{2}H and δ^{18}O measurements of the depleted samples (M and UM) from those values as back-calculated from the isotope data of mixtures (M10, M20, M50) and reference values of dilution standard std10 (Table 1 - Table 3). No obvious linearity problems are evident as would be visible by a significant difference/shift of values for direct measurements versus those back- calculated from dilutions.

Unfortunately, for the Picarro instrument at that time a temporary problem with a defective inlet valve (just recognized at the time of experiments) affected these measurements (especially visible in δ^{2}H results). No samples were left to perform the measurements again after replacement of the valve.

However, in general the results of Figure 4 show clearly that reliable isotope ratio measurements can be performed using laser techniques down to virtually δ^{2}H free samples and for δ^{18}O values down to at least -900 ‰ without any visible linearity problem. The best calculated estimate of the δ^{2}H composition of the depleted samples is at about -998‰.

5.3 Long-term reproducibility:

A newly established water standard (GRESP) was measured extensively for initial characterization over a period of 14 months at 19 individual measurement days with overall 1200 injections grouped in 52 sample mean values. The standard deviation of all mean values as shown in Figure 5a covering the whole time period is 0.08 ‰ for δ^{18}O and 0.7 ‰ for δ^{2}H. Daily calculated standard deviations for individual mean values are smaller at 0.06 ‰ and 0.4 ‰, respectively (see also [7]).

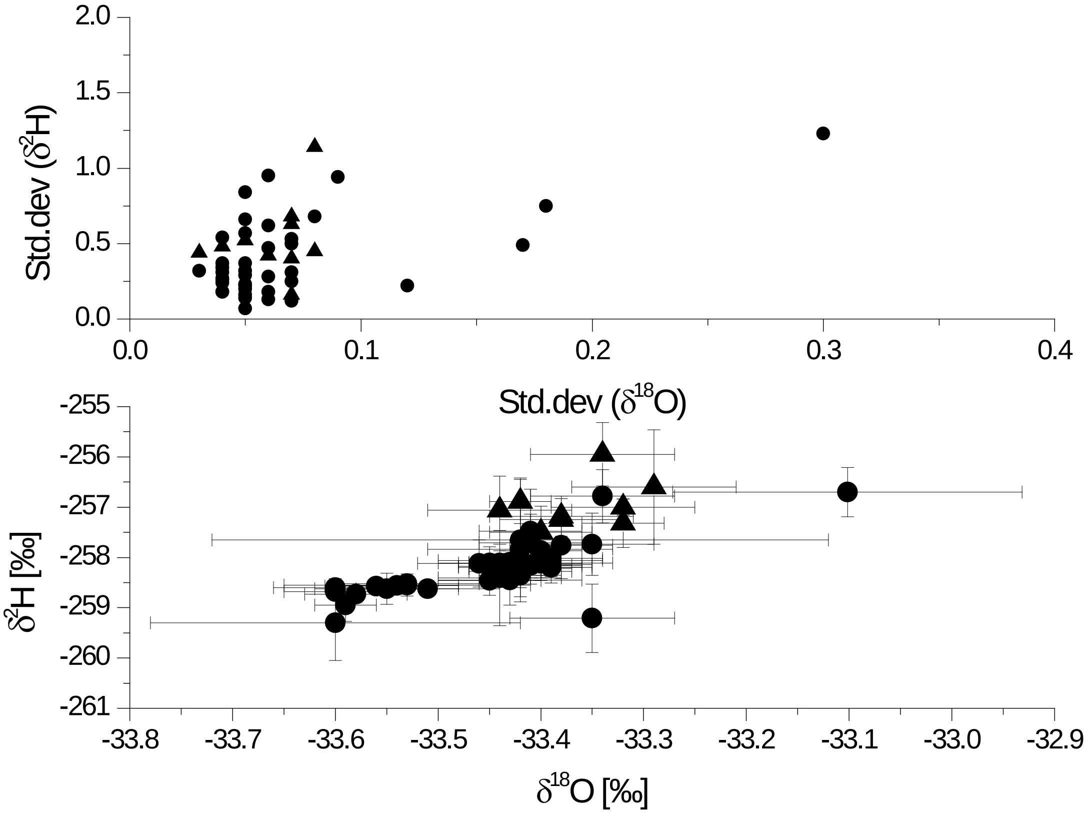

Figure 5: Long term assessment of a new standard and its standard deviation over a period of over one year with altogether 1200 injections performed.

5.4 Calibration exercise Internal Laboratory Standards:

Different analytical techniques were compared during two routine calibration exercises in March 2009 and March 2010, during which seven internal laboratory standards (ILS) were calibrated versus the international measurement standards VSMOW2 and SLAP2. Due to the previously already well known ILS isotopic values established in calibrations over the last ten years, an evaluation is possible concerning bias or accuracy of isotope data in addition to precision or standard deviation of measurements.

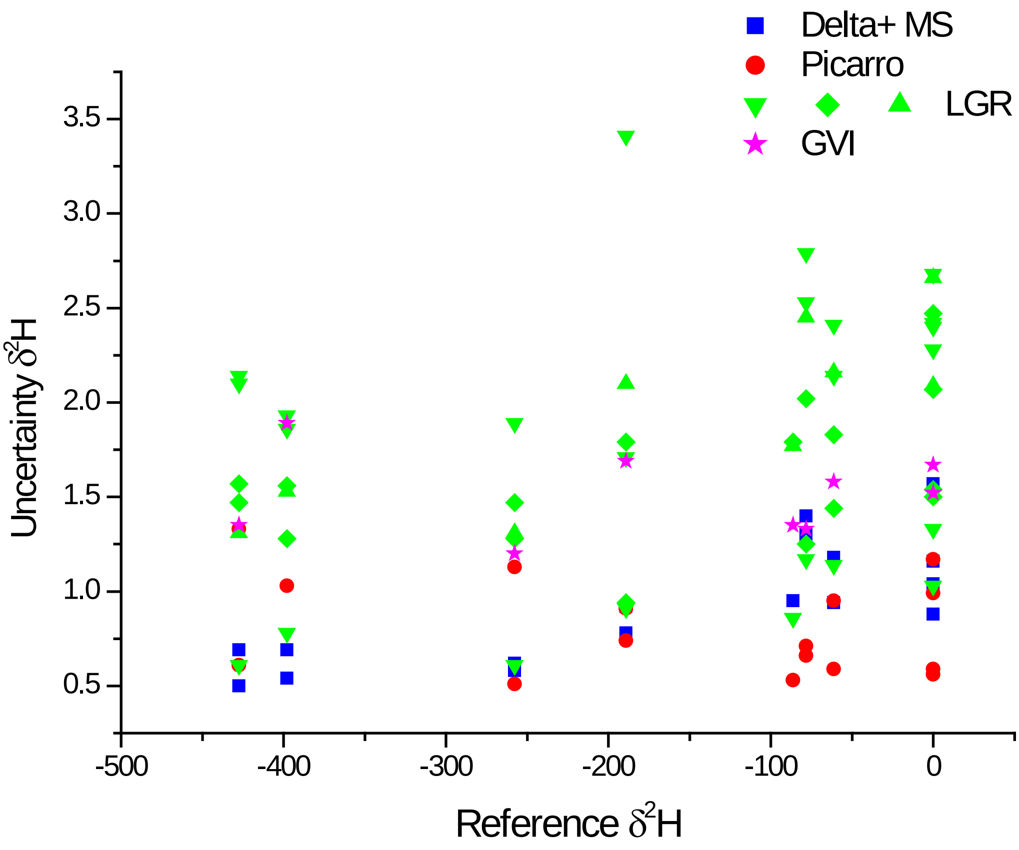

Figure 6: Combined standard uncertainties (1σ level) for δ^2H for a series of laboratory standards (indicated by same isotopic composition on x-axis) calibrated versus VSMOW2 and SLAP2, obtained during two calibrations performed in the years 2009 and 2010. The different symbols in this and in the following figures denote six different instruments used: square (Finnigan dual-inlet mass spectrometer Delta+); circle (Picarro laser); triangle up, triangle down and diamond (LGR laser); star (GVI continuous flow mass spectrometer Isoprime).

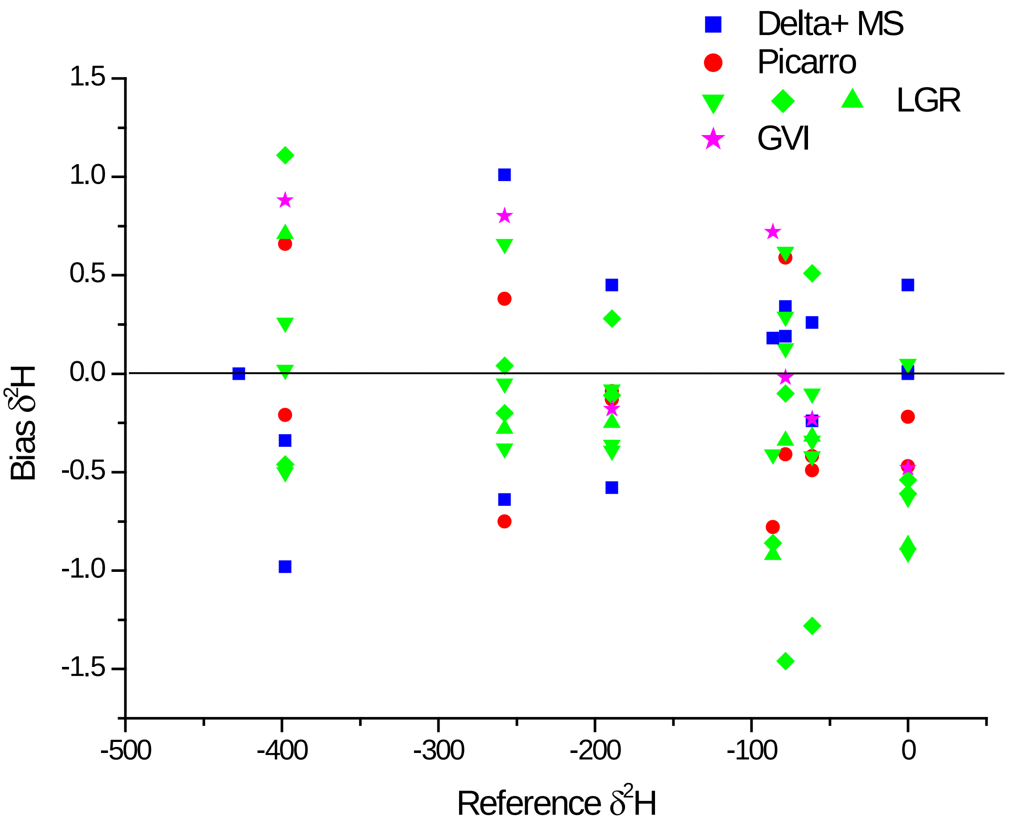

Figure 7: Calculated bias for performed calibrations versus previous reference values (for details of used symbols see Figure 6). Bias for VSMOW2 and SLAP2 used in the calibrations is zero per definition (squares at y-axis zero-line).

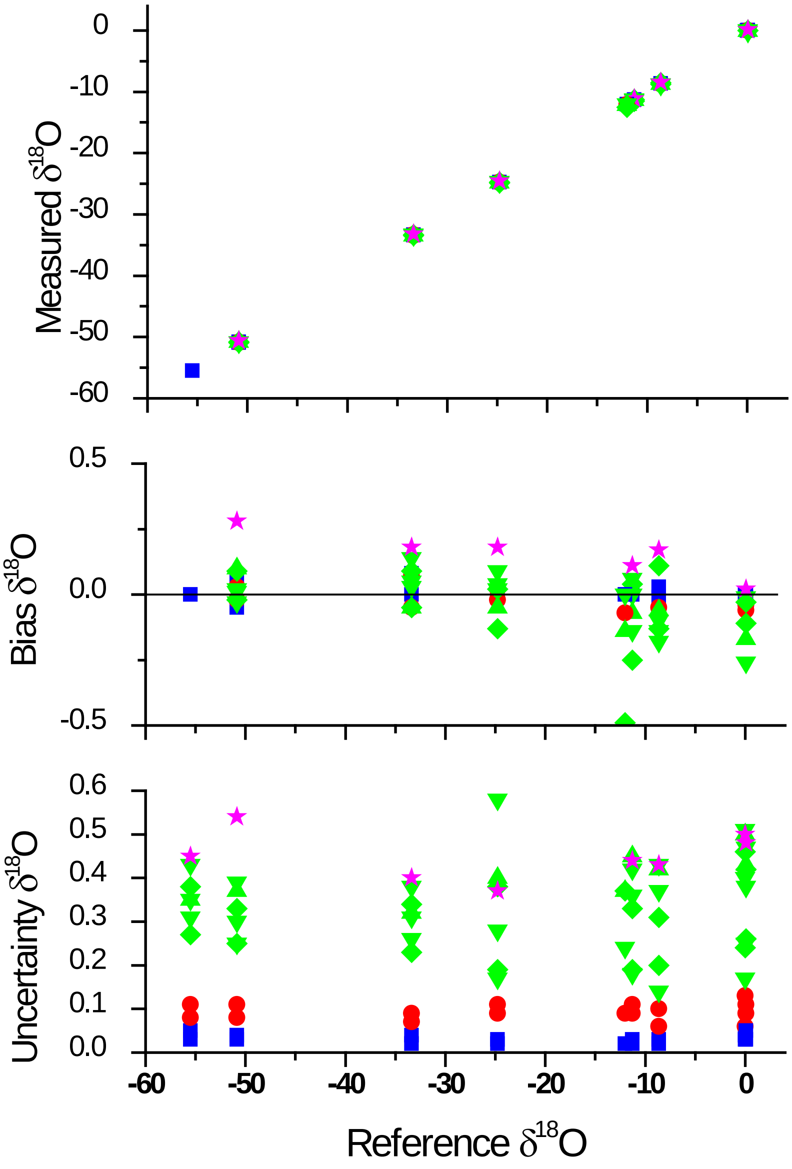

Figure 8: All data for $\delta^{18}O$ calibrations of laboratory standards as performed in the years 2009 and 2010 (for details of used symbols see Figure 6). Displayed are the measured values versus reference values (upper plot), the bias of measured values versus reference values for each instrument, and the combined standard uncertainty of mean values for each standard and instrument.

First it has to be noted that the results are all comparable within their combined uncertainties obtained by the different measurement techniques. The comparison included dual-inlet mass spectrometry with water/gas equilibration as the benchmark method (Finnigan Delta+ and self-built water/gas equilibration system). Further on until 2009 continuous flow mass spectrometry with high temperature pyrolysis system was used in the laboratory (GVI Isoprime with Hekatech HT-EA). In addition new infrared laser absorption techniques were applied in four instruments using two different measurement approaches (LGR off-axis ICOS, Picarro CRDS).

The achievable performance levels for the different instrumentation can be seen in Figure 6 and Figure 7 for $\delta^{2}H$, and in Figure 8 for $\delta^{18}O$ in its uncertainty graph (lower plot) and in the bias graph (middle plot, with a limit of below 0.2 ‰ for suitable analyses).

5.5 Calibration exercise External Laboratory Standards:

A calibration of external laboratory standards was performed as a follow up of a scientific meeting in December 2009 on water stable isotope analysis. Altogether 54 different water samples from 15 different laboratories worldwide used there as internal laboratory standards were analyzed at the IAEA IHL laboratory, involving three different methods: Dual-inlet mass spectrometry and two different infrared laser spectrometry approaches. Nearly 2500 injections / measurements were performed during that exercise. It was designed as calibration exercise, using IHL internal laboratory standards as calibrators. The findings and results were quite comparable to the results of the internal calibration exercise. Here the final results are provided as weighted mean values with weighted uncertainties. Due to the large number of repetitions the weighted uncertainties approach a level as small as 0.01 to 0.02 ‰ for $\delta^{18}O$ and 0.1 to 0.2 ‰ for $\delta^{2}H$. These values of course are just obtained as statistical averages and do not constitute an full uncertainty assessment. The average number of repetitions applied for the analytical instruments are around 2 for dual-inlet mass spectrometry, 12 for Picarro and 28 for LGR.

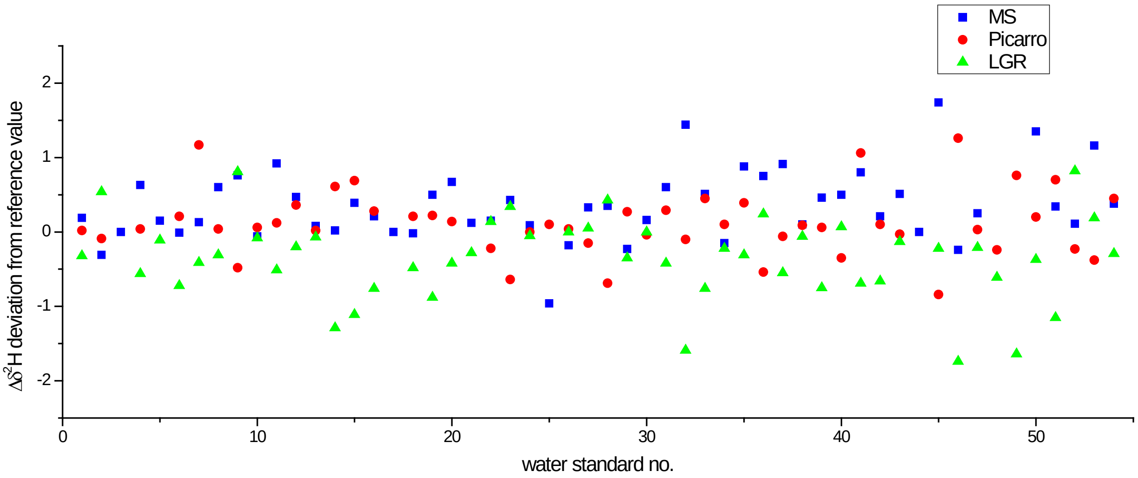

Figure 9: $\Delta\delta^2$H offsets from reference value for 54 external laboratory standards measured with three different methods at IHL. The reference value is the weighted mean of the three individual methods (also calculated as weighted means).

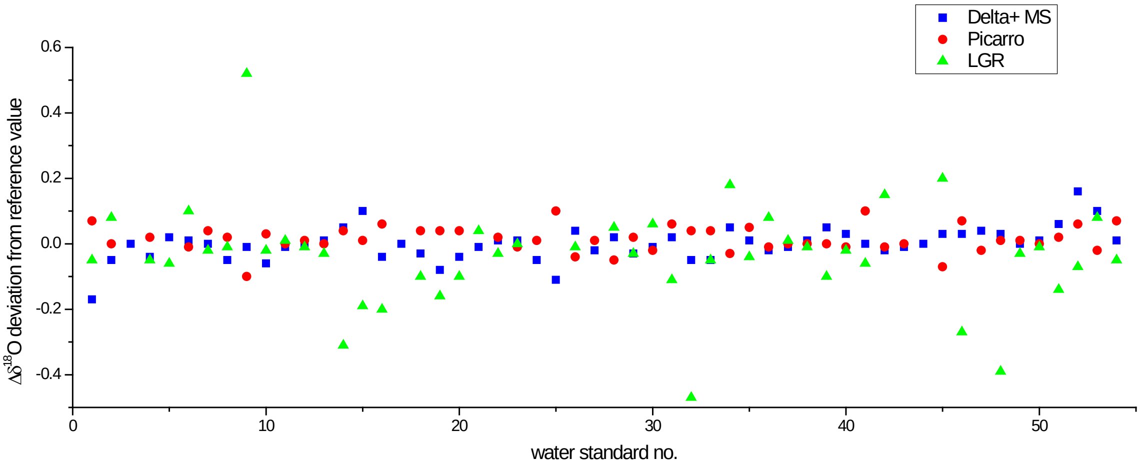

Figure 10: $\Delta\delta^{18}$O offsets from reference value for 54 external laboratory standards measured with three different methods at IHL. The reference value is the weighted mean of the three individual methods (also calculated as weighted means).

As general trend the overall virtual bias visible for both dual inlet mass spectrometry and Picarro CRDS laser spectrometry is well below 0.1 ‰ for $\delta^{18}O$ and at about 0.5 ‰ for $\delta^{2}H$ in accordance to other performed tests.

5.6 General reproducibility tests:

A series of reproducibility tests was performed both in April 2009 on a loaned Picarro system and again since October 2009 on a purchased system. The initial setup for the instruments used nitrogen gas from a high pressure gas cylinder. Routine settings included injection of 2 µL water, normally 6 times and sometimes 18 times per sample. After the first measurement days, the settings included 2 sample waste injections and 3 pre-flushes before actual sample injection. Later on the operation mode was switched from nitrogen gas to now use just laboratory air supplied by a small membrane pump through a Drierite-column at constant pressure (1200 hPa) as flushing gas into the instrument. This saves a lot of maintenance and costs for exchanging high pressure gas cylinders. No performance difference in isotopic results was noted for this change in dry gas supply. The daily standard deviation for measurements of known internal laboratory water standards was about ± 0.06 ‰ for $\delta^{18}O$ at an accuracy of better then ±0.05 ‰. For $\delta^{2}H$, daily standard deviations were in the order of 0.5 to 0.6 ‰, with accuracy values of ±0.6‰.

5.7 Decreased instrument performance:

For two months, problems occurred during measurement runs with a larger fraction of injections not producing any isotope result. Since no changes had been applied to the instrument, initially problems with the used syringe were proposed to be the reason. During further tests with different syringes the problems persisted. Sometimes for few hours measurements went quite normal, then one out of 15 injections failed and every about tenth injection showed higher water amounts. These water amount readings showed higher values up to 25000 ppm than the default 20000 ppm during a normal injection. Scrutinising the log files showed even larger water amounts for failing measurements (water amount signals exceeding 25000 ppm did not produce visible isotope data in the coordinator).

Still measurements performed during that time showed a similar performance than the other laser instrument type available (comparison with three LGR instruments and with Finnigan Delta+ MS, and later on comparison for isotopically depleted samples).

Finally with the help of a Picarro supplied valve diagnosis program a defective valve at the inlet to the evaporator chamber could be identified as cause of the problems. After exchange of the evaporator chamber with a new one the performance of the system was not compromised anymore.

Of course problems with clogged syringes occur during the course of measurements. For about 800 injections one is still on the safe side, afterwards at any time a syringe can show degraded performance, as visible in Figure 11 after about 30 injections. However, the instrument is very tolerant versus variances in injected water amount for individual measurements. While under normal conditions the injected water amount is constant within ±2 %, under unfortunate circumstances, like use of old syringes, the deviations in water amount can easily range towards ±25 %. Still the $\delta^{18}O$ values are deviating only by less than 0.2 ‰ and those of $\delta^{2}H$ by less than 1 ‰ (see Figure 11).

After change of the syringe, the performance is back to normal (shown in Figure 12 for $\delta^{2}H$).

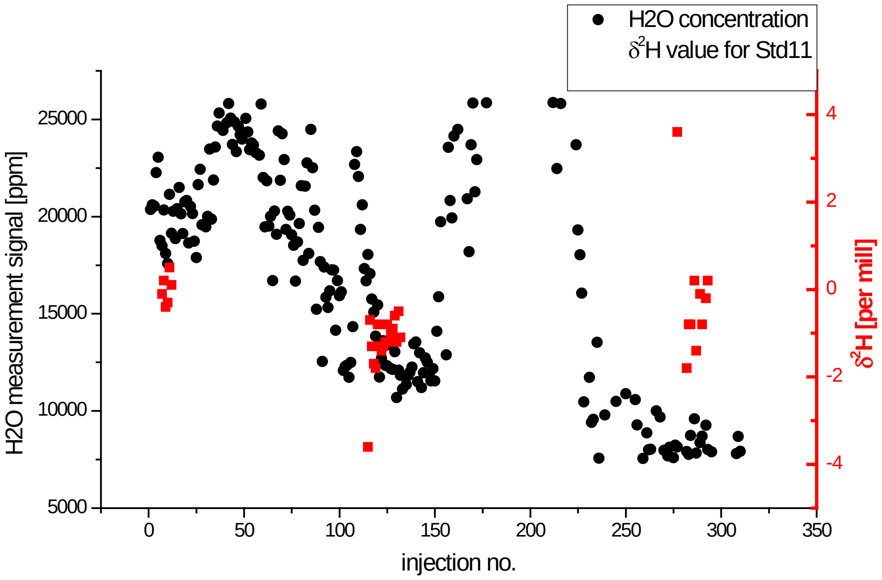

Figure 11: Injected water amount over time and resulting δ^2H values for laboratory Std11 (reference value -0.1±0.3‰) for a run with major syringe injection problems (2010-02-19). Still remarkable good reproducibility and performance for δ^2H measurements.

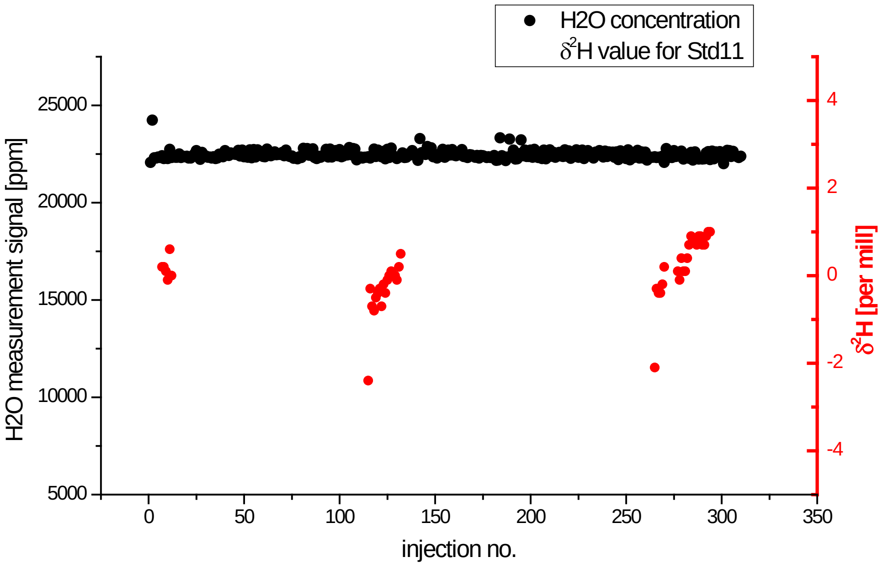

Figure 12: Injected water amount over time and resulting δ^2H values for laboratory Std11 for a normal run (2010-02-23).

6. Conclusions

The tested CRDS laser system (Picarro L1102-i) shows a high performance for groundwater measurements over a large range of isotopic compositions. Its performance for oxygen isotopes is similar to that of dual-inlet mass spectrometric systems using the water/gas equilibration technique and somewhat outperforms the mass spectrometer performance for hydrogen. Application of a robust external memory correction procedure ensures reliable isotopic results even under unfortunate measurement conditions. Its usability for isotopically largely enriched and depleted water samples is certainly interesting for applications much beyond the stable isotopic composition of natural water samples.

Acknowledgment

The supply of isotopically depleted water by Neil Sturcio, University of Illinois at Chicago, USA, is kindly acknowledged and very much appreciated, as it allowed to perform these measurements. Information received from a discussion with Bruce Vaughn, INSTAAR at University of Colorado, USA, triggered the development of the advanced memory correction used for result evaluations.

References

[1] BRAND, W.A., GEILMANN, H., CROSSON, E.R., RELLA, C.W., Cavity ring-down spectroscopy versus high-temperature conversion isotope ratio mass spectrometry; a case study on δ^2H and δ^{18}O of pure water samples and alcohol/water mixtures. Rapid Communications in Mass Spectrometry **23**(12) (2009) 1879-1884.

[2] LIS, G., WASSENAAR, L.I., HENDRY, M.J., High-Precision Laser Spectroscopy D/H and 18O/16O Measurements of Microliter Natural Water Samples. Analytical Chemistry **80**(1) (2008) 287-293.

[3] PICARRO. No compromise high throughput δ^{18}O and δD measurements (AN021). Picarro Application Notes 2009 [cited; Available from: http://www.picarro.com/resources/application_notes.

[4] GRÖNING, M., Improved water δ^2H and δ^{18}O calibration and calculation of measurement uncertainty using a simple software tool. Rapid Communications in Mass Spectrometry 2011, 25: 2711-2720, doi: 10.1002/rcm.5074

[5] SUCKOW, A. DUMKE, I., A database system for geochemical, isotope hydrological and geochronological laboratories. Radiocarbon 2001, 43, 325-337

[6] GRÖNING, M., SICalib User Manual (Stable Isotope Calibration for routine δ-scale measurements, Ver 2.16j). TEL Technical Note 01. IAEA Terrestrial Environment Laboratory, Vienna (2018) 37pp. link to be included here

[7] PICARRO. Exceptional Reproducibility from a Picarro Isotopic Water Analyzer (AN020). Picarro Application Notes 2008 [cited; Available from: http://www.picarro.com/resources/application_notes.