

On the reproducibility and repeatability of laser absorption spectroscopy measurements for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ isotopic analysis

D. Penna¹, B. Stenni², M. Šanda³, S. Wrede^{4,5}, T. A. Bogaard⁵, A. Gobbi¹, M. Borga¹, B. M. C. Fischer⁶, M. Bonazza², and Z. Chárová³

¹Department of Land and Agro-Forest Environments, University of Padova, Padova, Italy

²Department of Geosciences, University of Trieste, Trieste, Italy

³Faculty of Civil Engineering, Czech Technical University in Prague, Prague, Czech Republic

⁴Department Environment and Agro-Biotechnologies, Centre de Recherche Public – Gabriel Lippmann, Luxembourg

⁵Faculty of Civil Engineering and Geosciences, Delft University of Technology, Delft, The Netherlands

⁶Department of Geography, University of Zürich, Zürich, Switzerland

Received: 12 April 2010 – Published in Hydrol. Earth Syst. Sci. Discuss.: 12 May 2010

Revised: 26 July 2010 – Accepted: 30 July 2010 – Published: 13 August 2010

Abstract. The aim of this study was to analyse the reproducibility of off-axis integrated cavity output spectroscopy (OA-ICOS)-derived $\delta^2\text{H}$ and $\delta^{18}\text{O}$ measurements on a set of 35 water samples by comparing the performance of four laser spectrometers with the performance of a conventional mass spectrometer under typical laboratory conditions. All samples were analysed using three different schemes of standard/sample combinations and related data processing to assess the improvement of results compared with mass spectrometry. The repeatability of the four OA-ICOS instruments was further investigated by multiple analyses of a sample subset to evaluate the stability of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ measurements.

Results demonstrated an overall agreement between OA-ICOS-based and mass spectrometry-based measurements for the entire dataset. However, a certain degree of variability existed in precision and accuracy between the four instruments. There was no evident bias or systematic deviations from the mass spectrometer values, but random errors, which were apparently not related to external factors, significantly affected the final results. Our investigation revealed that analytical precision ranged from $\pm 0.56\text{‰}$ to $\pm 1.80\text{‰}$ for $\delta^2\text{H}$ and from $\pm 0.10\text{‰}$ to $\pm 0.27\text{‰}$ for $\delta^{18}\text{O}$ measurements, with a marked variability among the four instruments. The overall capability of laser instruments to reproduce stable results with repeated measurements of the same sample was acceptable, and there were general differences within the range of

the analytical precision for each spectroscopy. Hence, averaging the measurements of three identical samples led to a higher degree of accuracy and eliminated the potential for random deviations.

1 Introduction

In the past few decades, hydrogen and/or oxygen isotopes have been utilized in studies of different environments to address several areas of research in catchment hydrology, which include runoff generation processes (Brown et al., 1999; Weiler et al., 2003; Tetzlaff et al., 2007), preferential flow paths (Rodgers et al., 2005a,b; Lee et al., 2007; La Bolle et al., 2008), catchment and hillslope residence and transit time (McGuire and McDonnell, 2006; Lyon et al., 2008; Stewart et al., 2010), the contribution of pre-event and event water to the total stormflow (Uhlenbrook and Hoeg, 2003; Huth et al., 2004; Lyon et al., 2009), and the contribution of snowmelt in hydrograph separation applications (Taylor et al., 2002; Koeniger et al., 2008; Zhou et al., 2008).

The conventional method used to determine $\delta^{18}\text{O}$ and $\delta^2\text{H}$ (VSMOW-SLAP scale) in water samples is mass spectrometry (isotope-ratio mass spectrometry or IRMS). The disadvantages of this methodology are the time- and labour-intensive measurements coupled with the high equipment and operational costs. Recently, alternative instruments for isotopic analyses have been developed to offer more cost-effective opportunities for the determination of stable isotope ratios in the vapour or liquid water phase. Off-axis integrated cavity output spectroscopy (OA-ICOS) exploits Beer-



Correspondence to: D. Penna
(daniele.penna@unipd.it)

Lambert's law (Ricci et al., 1994) to relate the absorption of a laser light passing through a vaporized water sample to the isotopic composition of the sample. Therefore, OA-ICOS instruments allow for the simultaneous analysis of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ for each injection of water, reducing time and operational expenses per measured sample. In addition, simultaneous measurements exclude the potential relative error of two separate measurements of hydrogen and oxygen isotopes at different times. Further advantages include the reduced sample size (1–1.5 ml), easier maintenance requirements without extensive sample pre-processing, shorter time to produce reportable data, and the opportunity for in situ measurements in the field (Berman et al., 2009).

Recent studies (Aggarwal et al., 2006; Lis et al., 2008; Wassenaar et al., 2008; IAEA, 2009b; Singleton et al., 2009) have investigated the accuracy and reliability of laser absorption spectroscopy measurements of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ from water samples, and have underlined the main advantages of this technology. For instance, Lis et al. (2008) conducted a detailed investigation on the performance of the OA-ICOS analyzer and assessed the instrument precision, estimates of inter-sample memory and sample mass effect, and instrumental drift by comparing OA-ICOS-derived isotopic values with known standards. However, these studies only focused on the overall performance of a single machine or, in the case of measurements carried out by multiple analyzers (IAEA 2009b), the consistency of measurements among different instruments was not investigated. Moreover, shortcomings remained in the comparison of standardized schemes and analysis procedures in relation to traditional IRMS. Despite the breadth of available literature on the reliability and efficiency of laser spectroscopy, an inter-comparison test among various OA-ICOS analyzers over a significant number of water samples and under typical laboratory conditions was still absent. Therefore, the present work aimed to assess the following: (i) the reproducibility of measurements for four liquid water isotope analyzers over a 35 sample dataset; (ii) the overall performance of the four machines compared with a traditional mass spectrometer; (iii) the repeatability of each analyzer, i.e., the ability to constantly reproduce the same isotopic values; and (iv) the potential improvement in accuracy derived from the application of different analytic schemes and data-processing methods.

2 Materials and methods

2.1 OA-ICOS isotope analyzer and IRMS

All isotopic analyses were conducted using the off-axis integrated cavity output spectroscopy method with four liquid water isotope analyzers (LWIA), model DLT-100, which included three units version 908-0008 and one upgraded version 908-0008-2000 manufactured by Los Gatos Research Inc. (LGR, Mountain View, California, USA). Isotopic anal-

yses were performed at the Department of Land and Agro-Forest Environments at the University of Padova in Italy; the Faculty of Civil Engineering at the Czech Technical University in Prague; the Department of Environment and Agrobiotechnologies, Centre de Recherche Public - Gabriel Lippmann in Luxembourg; and the Faculty of Civil Engineering and Geosciences at the Delft University of Technology in the Netherlands. Each of these four analyzers was connected to a LC PAL liquid auto-injector (908-0008-9001, CTC Analytics AG, Zwingen, Switzerland) for the automatic and simultaneous measurement of $^2\text{H}/\text{H}$ and $^{18}\text{O}/^{16}\text{O}$ ratios in water samples. The auto-injector was provided with a 1.2 μl syringe (model 26P/-mm/AS, 7701.2N CTC) manufactured by Hamilton Company (Reno, Nevada, USA) for the injection of water samples into a heated port. All water samples and working standards were injected into ND8 32·11.6 mm screw neck 1.5 ml vials with PTFE/silicone/PTFE septums. The vials were filled with 1 ml of water and placed into 54 position trays on the auto-injector tray holder.

According to the manufacturer's specifications (Los Gatos Research Inc., 2008), the DLT-100 908-0008 LWIA provides isotopic measurements with a 1- σ precision below 0.6‰ for $\delta^2\text{H}$ and 0.2‰ for $\delta^{18}\text{O}$. The four analyzers in this comparative test were named I, II, III, and IV. Instrument IV refers to the upgraded model. Further information regarding the OA-ICOS theory of operation is reported in Paul et al. (2001); Baer et al. (2002); Sayres et al. (2009), and Wang et al. (2009).

Mass spectrometry analysis of the water samples was performed at the Isotope Geochemistry Laboratory of the Department of Geosciences, University of Trieste in Italy. Oxygen and hydrogen isotope measurements have been performed using the CO_2/H_2 water equilibration technique (Epstein and Mayeda, 1953; Horita et al., 1989). The equilibration device used for these analyses was a GFL 1086 connected to a Thermo Fischer Delta Plus Advantage mass spectrometer (Thermo Fisher Scientific Inc., Massachusetts, USA). The precision of $\delta^{18}\text{O}$ and $\delta^2\text{H}$ measurements, achieved with the IRMS technique was $\pm 0.05\%$ and $\pm 0.7\%$, respectively. Further information regarding the IRMS theory of operation and method is available in Roether (1970); Rolston et al. (1976); Hut (1987), and Horita and Kendall (2004).

2.2 Samples

Comparative analyses were performed on a dataset of 35 water samples (Table 1) characterized by a wide range of isotopic ratios. Isotopic contents ranged from -425% to -11% for $\delta^2\text{H}$ and from -55% to -1% for $\delta^{18}\text{O}$. Most of the samples originating from central-southern European streams, glaciers, rainfall, and snow exhibited intermediate isotopic contents ranging from -45% to -100% for $\delta^2\text{H}$ and from -5% to -15% for $\delta^{18}\text{O}$. Isotopic ratios of all samples measured by mass spectrometry were derived by

Table 1. Geographical origin and isotopic values (analyzed by mass spectrometry) of the study samples and reference standards. IT: Italy, CR: Czech Republic, USA: United States of America. Repeated samples are marked in bold.

ID	origin	$\delta^2\text{H}$ (‰)	standard deviation $\delta^2\text{H}$ (‰)	$\delta^{18}\text{O}$ (‰)	standard deviation $\delta^{18}\text{O}$ (‰)	standards used for analysis
1	tap water, Venice, IT	-59.71	0.32	-8.87	0.01	LGR2,3,4
2	Brenta river, Po plain, IT	-60.79	0.38	-9.03	0.02	LGR2,3,4
3	well, Po plain, IT	-44.07	0.37	-7.01	0.02	LGR3,4,5
4	alpine stream, IT	-85.11	0.30	-12.10	0.01	LGR2,3,4
5	alpine creek, IT	-79.73	0.45	-11.22	0.01	LGR2,3,4
6	alpine stream, IT	-80.28	0.18	-11.43	0.02	LGR2,3,4
7	alpine fountain, IT	-85.16	0.29	-11.88	0.01	LGR2,3,4
8	alpine creek, IT	-79.48	0.51	-11.42	0.01	LGR2,3,4
9	alpine stream, IT	-77.68	0.43	-11.16	0.02	LGR2,3,4
10	alpine stream, IT	-71.11	0.16	-10.40	0.01	LGR2,3,4
11	pre-alpine river, IT	-67.78	0.45	-9.89	0.01	LGR2,3,4
12	pre-alpine stream, IT	-71.14	0.49	-10.35	0.01	LGR2,3,4
13	mountain hail, IT	-61.07	0.41	-9.34	0.02	LGR2,3,4
14	alpine creek, IT	-59.80	0.44	-8.99	0.02	LGR3,4,5
15	alpine stream, IT	-89.15	0.17	-12.42	0.02	LGR2,3,4
16	alpine thermal spring, IT	-96.67	0.29	-13.28	0.02	LGR2,3,4
17	alpine stream, IT	-96.03	0.22	-13.59	0.02	LGR2,3,4
18	snow from alpine glacier, IT	-99.01	0.46	-13.32	0.02	LGR2,3,4
19	alpine stream, IT	-99.46	0.58	-13.90	0.02	LGR2,3,4
20	alpine stream, IT	-84.79	0.26	-11.98	0.01	LGR2,3,4
21	alpine stream, IT	-61.58	0.27	-9.15	0.01	LGR3,4,5
22	tap water, Padova, IT	-56.06	0.34	-8.50	0.03	LGR2,3,4
23	rainfall, Prague, CR	-20.87	0.26	-3.84	0.02	LGR3,4,5
24	mountain rainfall, CR	-12.32	0.27	-2.04	0.02	LGR3,4,5
25	tap water, Prague, CR	-31.48	0.35	-0.71	0.04	LGR3,4,5
26	evaporated tap water, Prague, CR	-67.65	0.59	-9.25	0.02	LGR3,4,5
27	mountain snow, CR	-72.86	0.28	-11.90	0.02	LGR2,3,4
28	mountain snow, CR	-107.28	0.23	-15.09	0.01	LGR1,2,3
29	mountain groundwater, CR	-72.15	0.33	-10.49	0.02	LGR2,3,4
30	Antarctic snow	-399.10	0.40	-51.14	0.03	TS6,7,8
31	Antarctic snow	-313.94	0.25	-39.49	0.00	TS6,7,8
32	Antarctic snow	-424.23	0.42	-54.67	0.01	TS6,7,8
33	Antarctic snow	-305.65	0.16	-38.29	0.01	TS6,7,8
34	Antarctic snow	-361.98	0.33	-45.74	0.01	TS6,7,8
35	stream, Hawaii Islands, USA	-10.86	0.18	-3.12	0.02	LGR3,4,5
LGR1	reference standard	-154.1	1.00	-19.57	0.10	–
LGR2	reference standard	-117.0	1.00	-15.55	0.10	–
LGR3	reference standard	-79.0	1.00	-11.54	0.10	–
LGR4	reference standard	-43.6	1.00	-7.14	0.10	–
LGR5	reference standard	-9.8	1.00	-2.96	0.10	–
TS6	reference standard	-224.4	0.50	-28.44	0.05	–
TS7	reference standard	-314.9	0.50	-40.26	0.05	–
TS8	reference standard	-423.5	0.50	-53.95	0.05	–

averaging the results of five repetitions. During the same run of each laser analyzer, seven samples were measured three times (marked in bold in Table 1), placed on the tray in three distinct vials, and treated as different samples, which generated 49 measurements for statistical analysis. A set of eight reference standards was used for all laser spectroscopy

analyses. The manufacturer provided five standards (named LGR1-5 in Table 1), that were intended for initial testing purposes during instrument installation. The origin, stability and characterization of their isotopic composition were not specified. However, as these standards were used for all analyses in this study, a potential systematic deviation from the

isotopic composition will not affect the comparative analyses among the four LWIAs. Issues could arise when comparing laser-derived with mass spectrometer-derived measurements but the excellent agreement between OA-ICOS-based and IRMS-based delta values (see Sect. 3.1) confirmed the declared isotopic composition of LGR reference standards. Therefore, the use of the five standards provided by the manufacturer was adopted in order to test the performance and the consistency of different analyzers in their standard setting as an average end user. For the extremely light Antarctic samples, three very negative standards were provided by the Isotope Geochemistry Laboratory of the University of Trieste, Italy. All standards were calibrated against IAEA (International Atomic Energy Agency) water standards (Gonfiantini, 1978) in relation to the VSMOW-SLAP scale and normalized adopting the procedure described in IAEA (2009a). The analysis required three reference standards that were bracketed around (i.e., slightly wider than) the isotopic composition of the unknown samples to determine their isotopic values. Therefore, samples were grouped according to their estimated isotopic ratio, and three appropriate standards were used (Table 1). Sample preparation, vial filling with disposable pipette tips, and labelling operations were executed in the laboratory of Padova to ensure consistency and homogeneity throughout the comparison.

2.3 OA-ICOS analysis schemes

To assess potential differences, all samples were evaluated with the following three analytic schemes (Fig. 1):

(i) Scheme (A) was proposed by the Isotope Hydrology Laboratory at IAEA (IAEA, 2009b). This procedure adopted two calibration standards and a control standard with an intermediate isotopic composition. Measurements and known δ values for calibration standards were interpolated by means of a linear regression to convert measured absolute $^2\text{H}/^1\text{H}$ and $^{18}\text{O}/^{16}\text{O}$ ratios to delta values. The control standard was not included in the calibration and therefore it could be used as indicator of the analysis accuracy by comparing the known value to the value measured by the laser spectroscope during the analysis. According to this scheme, each vial was sampled six times, and the first two measurements were discarded to reduce the memory effect (i.e., the influence of the previously injected sample on the isotopic content). Therefore, the reportable value was based on the average of the last four injections. Every run began with a dummy sample to prime the flow line and stabilize the system, and the last vial was filled with deionized water to clean the syringe (IAEA, 2009b). Standards were grouped in triplets, and samples formed sets of five unknowns.

(ii) Scheme (B) involved a calibration equation based on the interpolation of three standards. The scheme began with 36 injections of deionized water to clean the syringe and allow for the machine to warm up to operational temperature. Afterwards, each sample was injected six times, the

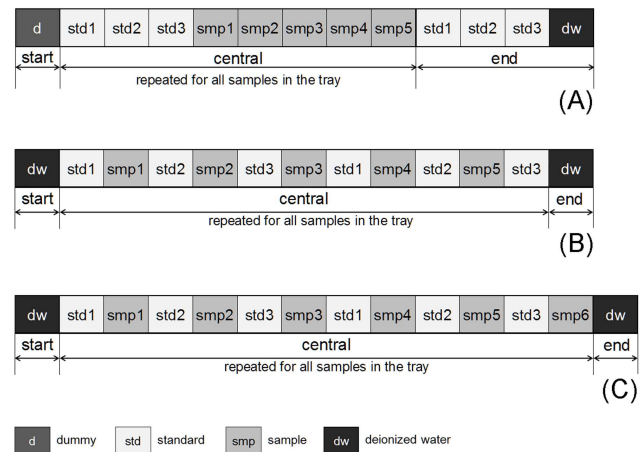


Fig. 1. Schematic representation of vial arrangement in the tray according to the different analytic schemes.

first measurement was rejected, and a preliminary mean was computed among the remaining five values. To avoid the influence of potential outliers, two measurements with the highest deviation from the preliminary mean were discarded, and the remaining three injections were averaged to obtain the reportable isotopic ratio. The run ended with 12 injections of deionized water from the first vial.

(iii) Scheme (C) was a modification of scheme (B), but each vial was injected eight times instead of six. The first three measurements were discarded. The reportable delta value was then obtained by averaging the three remaining injections, while two measurements with the highest deviations from the preliminary mean were discarded.

To determine the potential influence of different methods for averaging injections on the final isotopic values, all raw data were processed using the following three approaches: (i) the mean among the last four measurements of the six injections (or eight injections in scheme (C)) was referred to as version 1; (ii) the mean among the “best” three injections out of the last five was referred to as version 2; and (iii) the mean among six measurements after discarding the first two measurements in the case of (C) was referred to as version 3. The transfer line and syringe were cleaned at the start of each run to ensure that the inter-laboratory experimental conditions were as homogeneous as possible. A new heater septum, clean and dry vials with new cap septa, a new pipette tip for each sample or standard, and new or regenerated desiccants were used for every run. All samples and standards, which were usually stored at 4 °C, were kept at laboratory temperature for a minimum of 12 h and shaken to re-equilibrate the original isotopic composition prior to any analysis. On average, the cavity operational temperature of the four analyzers for the comparative runs ranged between 26 and 29 °C.

2.4 Statistical tests

To assess the performance of the LGR laser analyzers compared with a mass spectrometer, the deviations between the OA-ICOS-derived and IRMS-derived measurements were computed for the whole dataset:

$$\Delta_H(\%) = \delta^2 H_{\text{OA-ICOS}} - \delta^2 H_{\text{IRMS}} \quad (1)$$

$$\Delta_O(\%) = \delta^{18} O_{\text{OA-ICOS}} - \delta^{18} O_{\text{IRMS}} \quad (2)$$

where $\delta^2 H_{\text{OA-ICOS}}$ and $\delta^{18} O_{\text{OA-ICOS}}$ were the isotopic delta values determined by the laser spectroscope for the hydrogen and oxygen isotopes, respectively, while $\delta^2 H_{\text{IRMS}}$ and $\delta^{18} O_{\text{IRMS}}$ were the isotopic delta values determined by the mass spectrometer for the hydrogen and oxygen isotopes, respectively. Therefore, a perfect agreement between the laser spectroscope and the mass spectrometer measurements was achieved when $\Delta_{H,O} = 0$, and the laser spectroscope overestimated or underestimated the mass spectrometer values for $\Delta_{H,O} > 0$ and for $\Delta_{H,O} < 0$, respectively.

To assess the statistical significance of deviations between the OA-ICOS and the IRMS measurements, a one-sample t-test was performed to compare the mean of each deviation series to a hypothesized value equal to zero (i.e., no deviation present between spectroscopy and spectrometry measurements). A multifold approach to test the normality of each deviation series was followed. First, frequency histograms and normal probability plots (not reported herein) were utilized to visually assess the potential deviation of each distribution from the theoretical Gaussian curve. Second, Shapiro-Wilk, Kolmogorov-Smirnov, and Anderson-Darling normality tests were performed at a significance level of 0.05. The combined application of these three approaches reduced the possibility of rejecting a normal-distributed series as non-normal or vice versa. Since this preliminary analysis demonstrated the departure of several OA-ICOS - IRMS deviations from the Gaussian distribution, a non parametric approach was followed to statistically evaluate the differences between OA-ICOS and mass spectrometry-derived isotopic measurements. Thus, the one-sample t-test was performed for the normal distributed deviation series, whereas the one-sample sign test was applied to non-normal error distributions. Under the null hypothesis that no difference existed between the observed and assumed median of zero, the one-sample sign test considered that the probability of finding observations above the assumed median should be equal to the probability of finding observations below the assumed median. Therefore, the one-sample t-test involved the formulation of the following null and alternative hypotheses:

$$H_0 : \mu = \mu_0 \quad (3)$$

$$H_1 : \mu \neq \mu_0 \quad (4)$$

where μ was the mean of the laser-IRMS deviation series and μ_0 was the hypothesized mean (placed equal to zero).

For the one-sample sign test, the following null and alternative hypotheses were formulated:

$$H_0 : \eta = \eta_0 \quad (5)$$

$$H_1 : \eta \neq \eta_0 \quad (6)$$

where η was the median of the OA-ICOS - IRMS deviation series and η_0 was the hypothesized median (i.e., equal to zero).

3 Results and discussion

3.1 OA-ICOS - IRMS correlations and deviations

Scatterplots of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ allowed for a first comparison of laser- and mass spectrometry-based measurements for each of the four analyzers in this study. The plots (reported on <http://www.isotope-hydrology.net/>) exhibited excellent agreement between IRMS and OA-ICOS for both isotope species and for all analyzers, with determination coefficients (R^2) ranging from 0.99988 to 0.99996 for hydrogen and from 0.99929 to 0.99982 for oxygen ($n=49$). This observation confirmed previous results (Aggarwal et al., 2006; IAEA, 2009b; Singleton et al., 2009) analyzing different natural water samples. Despite the high values of the determination coefficients, there were slight variations between the four machines and two water isotopes, which indicated potential differences in instrumental behaviour. To assess the performance of laser spectroscopes with respect to the mass spectrometer, the distributions of the OA-ICOS - IRMS deviations were compared by the box-plots depicted in Figs. 2 and 3 for hydrogen and oxygen, respectively. For each of the four laser spectroscopes, three analytic schemes and averaging methods were applied. The plots suggested two main observations. Firstly, a certain degree of variability existed among the four machines for the two isotopic ratios, both in terms of accuracy (distance of the mean from the zero line) and precision (amplitude of the boxes, i.e., standard deviation). In particular, the distributions of errors for hydrogen analyses using machine II (Fig. 2, panel II) displayed lower standard deviations compared with other machines. However, the lines representing the mean always plotted above the zero line, which revealed a constant overestimate with respect to the IRMS. In contrast, all deviation series exhibited a relatively high accuracy for hydrogen isotopic measurements on instrument IV (Fig. 2, panel IV). For oxygen, analyzers I and IV (Fig. 3 panels I and IV, respectively) exhibited measurements that were slightly underestimated with respect to the mass spectrometer. Contrary to the performance for hydrogen, analyzer II generated relatively accurate measurements of oxygen (Fig. 3, panel II). Analyzer III exhibited the maximum span of deviation and the highest degree of variability (Fig. 3, panel III), which indicated a lack of precision compared with the other spectroscopes.

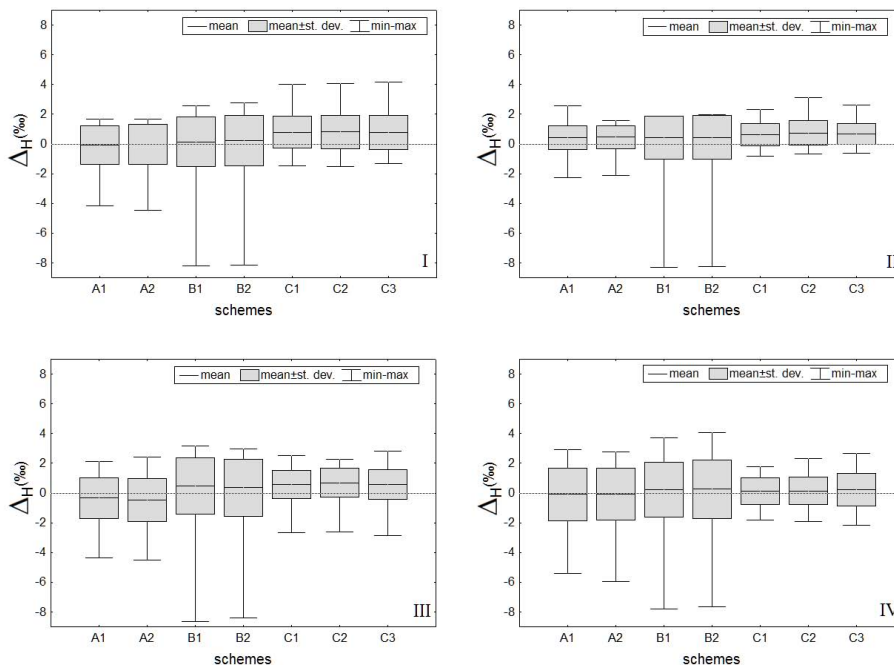


Fig. 2. Descriptive statistics of $\delta^2\text{H}$ laser spectroscopy-mass spectrometer deviations for the different analytic schemes and for the four analyzers.

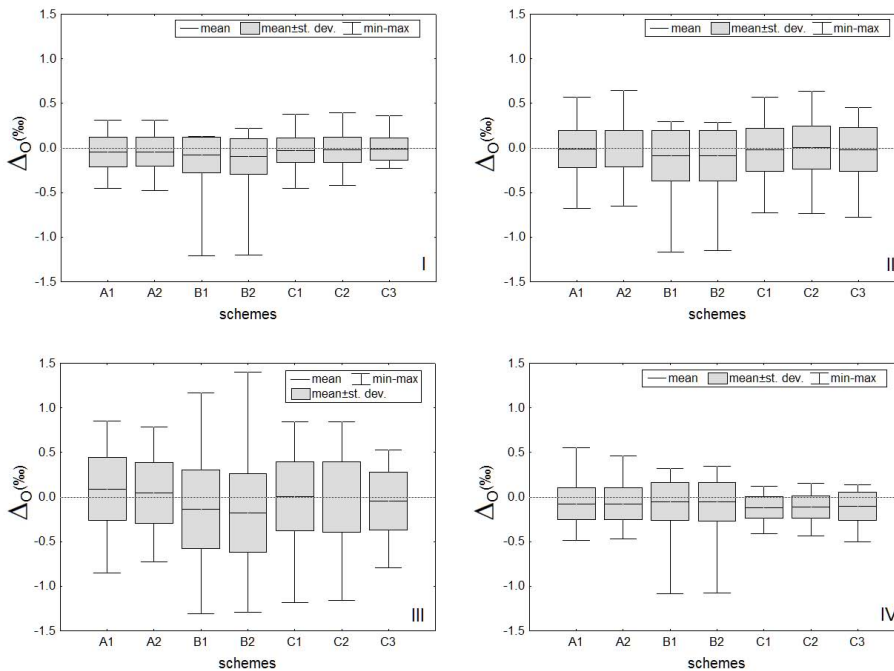


Fig. 3. Descriptive statistics of $\delta^{18}\text{O}$ laser spectroscopy-mass spectrometer deviations for the different analytic schemes and for the four analyzers.

Secondly, the use of different analytic schemes did not seem to be a determining factor in the improvement of isotopic measurements by laser spectroscopy. Indeed, no scheme was able to consistently provide the most accurate

and precise or worst measurements for all instruments and both isotopes. Moreover, no systematic behaviour (i.e., constant overestimation or underestimation, constant low or high standard deviation) was observed among the schemes, and

Table 2. Basic descriptive statistics for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ laser spectroscopy - mass spectrometer deviations for the different schemes and for the four analyzers (in bold is the mean value closest to zero for each isotope; in italics is the lowest standard deviation value for each isotope).

LWIA	scheme	$\delta^2\text{H}$ deviations (‰)		$\delta^{18}\text{O}$ deviations (‰)	
		mean	standard deviation	mean	standard deviation
I	A1	-0.09	1.30	-0.04	0.17
	A2	-0.03	1.34	-0.04	0.16
	B1	0.15	1.68	-0.08	0.20
	B2	0.22	1.71	-0.09	0.20
	C1	0.78	1.07	-0.03	0.14
	C2	0.80	1.12	-0.02	0.14
	C3	0.78	1.14	-0.01	0.12
II	A1	0.42	0.81	-0.01	0.21
	A2	0.45	0.79	-0.01	0.20
	B1	0.43	1.43	-0.09	0.28
	B2	0.45	1.46	-0.09	0.28
	C1	0.64	0.75	-0.02	0.24
	C2	0.75	0.84	0.01	0.24
	C3	0.66	0.70	-0.02	0.24
III	A1	-0.33	1.37	0.09	0.35
	A2	-0.49	1.45	0.05	0.34
	B1	0.47	1.90	-0.14	0.44
	B2	0.37	1.93	-0.18	0.44
	C1	0.55	0.95	0.01	0.39
	C2	0.69	0.99	0.00	0.40
	C3	0.57	1.00	-0.05	0.33
IV	A1	-0.08	1.76	-0.08	0.18
	A2	-0.08	1.74	-0.07	0.18
	B1	0.21	1.85	-0.05	0.21
	B2	0.26	1.97	-0.05	0.22
	C1	0.12	0.88	-0.12	0.12
	C2	0.15	0.91	-0.11	0.12
	C3	0.21	1.11	-0.10	0.16

the inter-machine variability seemed to exceed the scheme variability. Furthermore, almost no difference existed between the two (or three for scheme (C)) versions of the averaging methods, which always yielded similar deviation distributions for all machines and both isotopic ratios. These results suggested that the influence of including or excluding possible outliers (i.e., injections that deviate the greatest from the preliminary mean) is minimal. Generally, the measurement was not improved by discarding such outliers, but less robust results would be generated due to the lower number of measurements considered. All of these observations were confirmed by the data in Table 2 which reports mean and standard deviation values for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ OA-ICOS-IRMS error distributions from the different schemes and four analyzers.

3.2 Statistical significance of OA-ICOS-IRMS deviations and analytic schemes

The significance of the deviations in the laser spectroscopy-mass spectrometry was assessed by a t-test and sign test (Table 3). The difference between the OA-ICOS and IRMS-derived measurements was not statistically significant at 95% if the p -value was greater than the significance level of 0.05 for both tests. In such cases, the spectroscopy deviations from the mass spectrometer measurements were negligible. This condition was met for hydrogen under a few instances for three of the four spectroscopes, reflecting the variability among the different machines. Analyzers I and III only provided accurate measurements in comparison to mass spectrometry when scheme (A) was applied, whereas instrument II yielded measurements that were always significantly different from IRMS, which confirmed the overestimation (see also Fig. 2). In contrast, spectroscopy IV almost always produced results with insignificant deviations compared with mass spectrometry. The performance of the four analyzers was different when the isotopic measurements for oxygen were considered. Analyzers I and III did not yield significantly different values from IRMS for all schemes (with relatively high p -values), except for (B). Analyzer II always provided accurate results that were independent of the applied scheme, which contrasted with the hydrogen measurements. Spectroscopy IV exhibited reliable results only with scheme (B2).

In general, there was not an absolute best scheme for accurate and precise results. However, scheme (A1), which was the original approach first described by IAEA (2009b) and accounted for the most rapidly generated data, was the scheme that most often resulted in values that were not significantly different compared with IRMS. For these reasons, scheme (A1) was considered the most representative.

3.3 Inter-machine variability and relation to sample isotopic composition

Despite the use of the same dataset, analytic scheme, averaging method, and instrumentation (only localized in different laboratories), significant inter-machine variability among the four laser analyzers and IRMS was observed. The source of this variability remains unknown: the potential effect on the results of variable water molecule density per injection and the water vapour temperature in the cavity was assessed, but no clear relationship could be identified between these variables and measurement errors (results not presented herein). Furthermore, no causal relations of OA-ICOS-IRMS deviations with external factors could be determined to explain these occasional deviations, which agreed with the empirical observation that “each analyser has its own idiosyncrasies” (Newman, B. D., personal communication, 2009).

Table 3. Results of one-sample t-test and one-sample sign test for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ deviations. The analyses that produced negligible differences between laser spectroscopy and mass spectrometer (i.e., not significant difference with $\alpha:0.05$) are marked in bold.

		t-test for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ deviations				Sign test for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ deviations			
		$\delta^2\text{H}$		$\delta^{18}\text{O}$		$\delta^2\text{H}$		$\delta^{18}\text{O}$	
LWIA	scheme	t	p-value	t	p-value	median	p-value	median	p-value
I	A1	–	–	–1.84	0.07	0.19	0.77	–	–
	A2	–	–	–1.85	0.07	0.33	0.77	–	–
	B1	–	–	–	–	0.43	0.04	–0.04	0.02
	B2	–	–	–	–	0.47	0.01	–0.06	0.00
	C1	–	–	–1.41	0.16	0.63	0.00	–	–
	C2	4.99	0.00	–1.11	0.27	–	–	–	–
	C3	4.81	0.00	–0.69	0.49	–	–	–	–
II	A1	–	–	–0.49	0.63	0.44	0.00	–	–
	A2	–	–	–0.25	0.80	0.63	0.00	–	–
	B1	–	–	–	–	0.71	0.00	–0.03	0.08
	B2	–	–	–	–	0.77	0.00	–0.03	0.39
	C1	6.00	0.00	–0.57	0.57	–	–	–	–
	C2	6.23	0.00	0.16	0.87	–	–	–	–
	C3	6.54	0.00	–0.48	0.64	–	–	–	–
III	A1	–	–	1.81	0.08	–0.26	0.25	–	–
	A2	–2.37	0.02	0.94	0.35	–	–	–	–
	B1	–	–	–2.18	0.03	0.78	0.00	–	–
	B2	–	–	–2.80	0.01	0.65	0.00	–	–
	C1	–	–	0.16	0.88	0.72	0.00	–	–
	C2	–	–	–	–	0.89	0.00	0.03	0.77
	C3	4.00	0.00	–1.01	0.32	–	–	–	–
IV	A1	–	–	–	–	0.29	0.25	–0.10	0.00
	A2	–	–	–	–	0.37	0.08	–0.11	0.00
	B1	–	–	–	–	0.55	0.01	–0.03	0.02
	B2	–	–	–	–	0.61	0.08	–0.02	0.25
	C1	0.99	0.33	–6.83	0.00	–	–	–	–
	C2	1.17	0.25	–6.45	0.00	–	–	–	–
	C3	1.36	0.18	–	–	–	–	–0.06	0.04

Nevertheless, a certain degree of error was related to the extreme isotopic content of the analyzed samples. Figure 4 shows the deviations between OA-ICOS- and IRMS-derived measurements for the whole dataset (49 samples with seven repetitions included) for $\delta^2\text{H}$ (panel a) and $\delta^{18}\text{O}$ (panel b). Results plotted above or below the zero line (indicating the perfect agreement between the two measurement approaches) and exhibited no regular structure for any instrument or isotopic content. The only exception, especially for hydrogen, was the clear underestimated measurement performed by all laser spectroscopes for very light samples (more negative than -300‰ $\delta^2\text{H}$). This behaviour could be attributed to the memory effect which can have an important influence when analyzing extreme isotopic values (IAEA, 2009b). In such cases, discarding the first two injections and averaging the remaining four measurements could not be sufficient to overcome the problem. This effect was observed in approximately 50% of the light sample measurements that

were affected by an unknown source of error. Moreover, this behaviour was clearly marked for $\delta^2\text{H}$ readings, which revealed that the difference in accuracy could potentially affect these devices. Therefore, further investigations on this issue with special testing procedures are advised. Since no enriched samples over -10‰ $\delta^2\text{H}$ and -0.71‰ $\delta^{18}\text{O}$ were included in the dataset, predictions about potential similar behaviour in these data ranges were not possible. However, analyses executed by IAEA (2009b) on a set of artificial water samples up to approximately $+1670\text{‰}$ $\delta^2\text{H}$ and $+14\text{‰}$ $\delta^{18}\text{O}$ demonstrated comparable results to mass spectrometry, revealing a satisfactory throughput of laser analyzers at least on the positive side of the scale. In general, the analysis of samples with extremely positive or negative isotopic compositions by means of LGR laser spectroscopes should be performed carefully due to potential over- or underestimation errors.

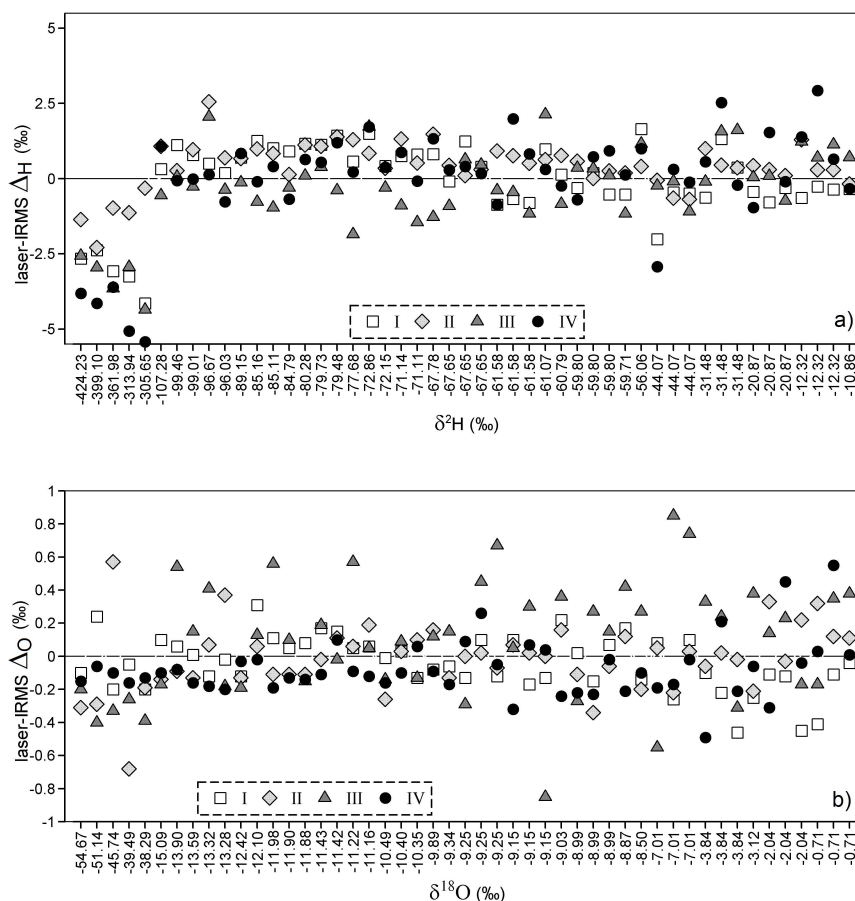


Fig. 4. $\delta^2\text{H}$ (a) and $\delta^{18}\text{O}$ (b) deviations for all datasets (A1 scheme). The horizontal line represents 0 value, i.e., no difference between laser spectroscopy and mass spectrometer measurements.

3.4 Performance examples

A few examples of the overall performance of the laser spectroscopes for both $\delta^2\text{H}$ and $\delta^{18}\text{O}$ measurements are reported in Fig. 5. The four panels (a–d) present comparisons between OA-ICOS- and IRMS-derived measurements for four samples featuring a different range of isotopic composition, and allow for the assessment of both accuracy (vicinity to the origin, where the mass spectrometry-derived value was placed) and precision (width of error bars, which reproduced the standard deviation of each measure) of OA-ICOS measurements for the two isotopic ratios and the four analyzers. Therefore, the following conclusions could be drawn from Fig. 5. First, the inter-machine variability was apparent in the different degrees of deviation for the mass spectrometer values across the four samples. However, no instrument exhibited the overall best or worst performance in terms of accuracy and precision. Second, biases or systematic deviations were not evident for any particular instrument with respect to the isotopic content of samples, except for the marked underestimation of all analyzers for measurements of very light

samples (Fig. 5d). The values determined by analyzers I, III, and IV grouped closely, while machine II exhibited isotopic measurements that were less underestimated because of a compensation effect from the usual positive deviation of the actual $\delta^2\text{H}$ value. Despite the lack of accuracy for hydrogen measurements of light samples, standard deviation values in the -300‰ to -400‰ $\delta^2\text{H}$ range were comparable or lower than those obtained by measurements of other samples, which suggested that instrumental precision was independent from the sample isotopic content. In general, no apparent relationship was identified between accuracy and precision of the two water isotopes. Thus, a good spectroscopy performance in measuring hydrogen isotopic content did not guarantee a similar performance for oxygen or vice versa.

3.5 Precision

Previous investigations have revealed different estimates for spectroscopy precision. Aggarwal et al. (2006) reported a degree of precision of $\pm 1\text{‰}$ for hydrogen and $\pm 0.3\text{‰}$ for

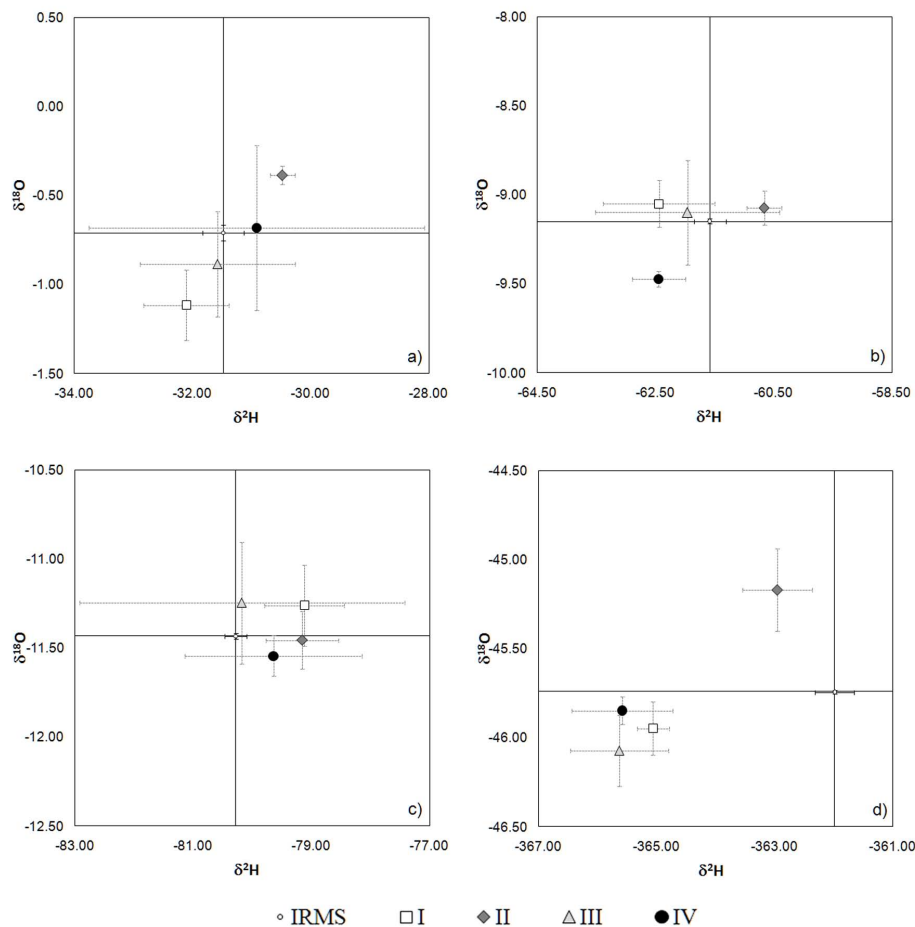


Fig. 5. Overall accuracy and precision of laser spectroscopy measurements compared with the mass spectrometer for a range of isotopic values (A1 scheme). Samples: (a) 29; (b) 25; (c) 8; (d) 38.

oxygen, using a prototype version of the LWIA featuring a different configuration than the commercial version. Lis et al. (2008) suggested a systematic sample analysis and data normalization procedure routine that resulted in a precision of $\pm 0.22\text{‰}$ for $\delta^2\text{H}$ and $\pm 0.16\text{‰}$ for $\delta^{18}\text{O}$. A few recent studies have indicated different values for 1- σ standard deviations. Researchers from the Isotope Hydrology Section of the IAEA determined a precision of approximately $\pm 1\text{‰}$ for $\delta^2\text{H}$ and $\pm 0.2\text{‰}$ for $\delta^{18}\text{O}$ (IAEA, 2009b), while Lyon et al. (2009) obtained precision values of $\pm 0.37\text{‰}$ for hydrogen and $\pm 0.12\text{‰}$ for oxygen, for the DLT-100, 908-0008 after measuring a reference standard with known isotopic content for more than six months. In our comparative analysis, we observed a marked difference in precision among the four spectroscopes. Table 4 presents the basic statistics of standard deviation values obtained in $\delta^2\text{H}$ and $\delta^{18}\text{O}$ measurements for the dataset of 49 samples. The variable behaviour of the four spectroscopes was evident when the statistical properties of the standard deviations were considered. For hydrogen, machine I performed the best in terms of preci-

sion with 75% of the measurements yielding a standard deviation less than $\pm 0.72\text{‰}$, which satisfies even complex hydrological applications of $\delta^2\text{H}$. In contrast, machines IV and III exhibited standard deviations in the hydrogen measurement that were noticeably different from the values reported in the literature, with means greater than $\pm 1\text{‰}$. The 25th and 75th percentiles suggest a lack of precision that, independently from machine accuracy, can affect the ability to analyze physical processes in the field, especially when differences in the water isotopic content are below 2‰ $\delta^2\text{H}$. A performance contrary to hydrogen was observed for oxygen (Table 4), with spectroscopy IV offering good precision. Except for instrument III, the analysers were characterised by a comparable or better precision than reported by the manufacturer (Los Gatos Research, Inc., 2008, 2010) or by the aforementioned studies. Spectroscopy III provided the highest standard deviation, and lacked precision (but not accuracy) for both water isotopes.

3.6 Precision and accuracy improvement

Our analyses did not suggest any evidence of factors that explained the variable behaviour of the four laser analyzers. Therefore, such differences were accounted for by white noise, which can be difficult to erase or reduce. During the post-processing phase, no “data cleaning” was performed and all raw data provided by the four instruments were reported to offer an undisturbed comparative view of the spectroscope performance. However, the results of each run must be observed carefully to detect possible “bad injections”, i.e., spikes or large dips in the amount of water sampled by the syringe. A number of water molecules in the range of $2\text{--}4 \times 10^{16}$ per cm^3 should be maintained during the run. If the number of molecules introduced into the laser cell cannot be stabilized in this expected range, the absorption peaks can be significantly influenced and higher uncertainties in the isotope ratios are likely to occur (Aggarwal et al., 2006; IAEA, 2009b). A dramatic case of this behaviour can be exemplified by Fig. 6, which reports the variations in water molecules injected into the cavity during a run performed by analyzer IV (scheme (A1)). The number of molecules per cm^3 was within the expected range and no noticeable trend or drift could be observed. Nevertheless, few injections were outside of the average pattern. The most prominent was a marked dip that occurred during injection number 257, which corresponded to the fifth of six for the determination of standard LGR3. Injection 257 yielded a water volume (3.13×10^{16} molecules/ cm^3) that was significantly less than the mean (3.48×10^{16} molecules/ cm^3) for the entire 270 injections of the run. This inconsistent water amount matched reportable delta values of hydrogen and oxygen, which were significantly different from the three values used for the final determination of the sample isotopic composition by the average of the last four injections (Table 5). The known isotopic content of standard LGR3 was -79.00‰ and -11.54‰ for $\delta^2\text{H}$ and $\delta^{18}\text{O}$, respectively. The average of the last four injections, which included number 257, provided a reportable delta values of -0.55‰ and -11.16‰ for $\delta^2\text{H}$ and $\delta^{18}\text{O}$, respectively. The deletion of injection 257 greatly reduced the standard deviation (from 5.43‰ to 1.69‰ for hydrogen and from 0.70‰ to 0.37‰ for oxygen) and improved the accuracy for $\delta^2\text{H}$ to provide a final value of -77.93‰ , which was closer to the known reference (injection number 257 included). Unfortunately, this operation did not improve the accuracy for $\delta^{18}\text{O}$ measurement to yield a final value of -10.85‰ , which was further from the actual value. This different behaviour can most likely be attributed to the general deviations of oxygen measurements from IRMS, which characterized spectroscope IV. These results clearly demonstrated the potential influence of inconsistent injections. No evident factor was responsible for such peculiar behaviour, but the intrinsic variability of the instrument. Moreover, not all injections that deviated from the average volume of water corresponded to inconsistent isotopic values. Nevertheless,

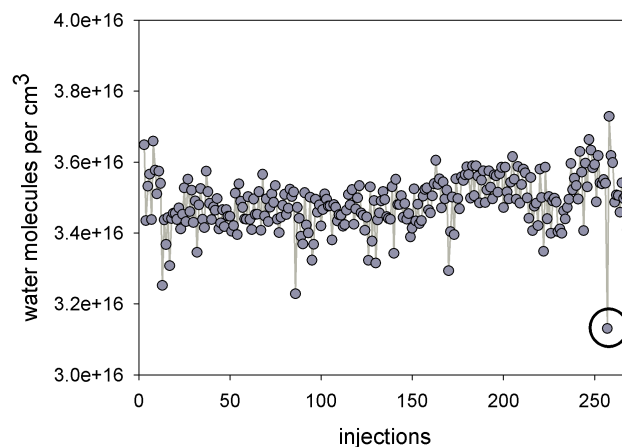


Fig. 6. Example of inconsistent injection.

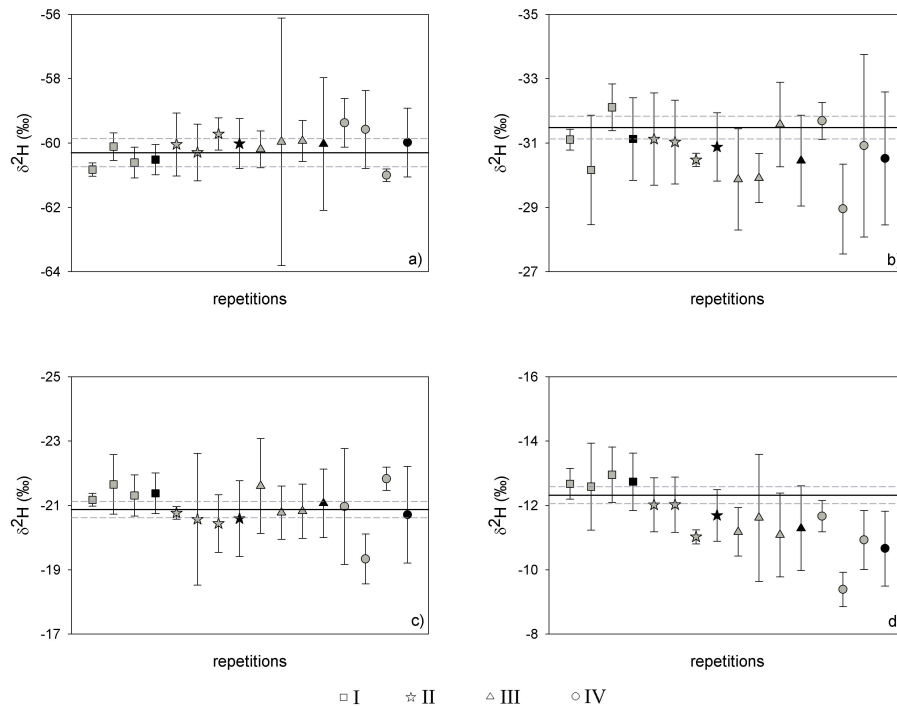
a close inspection of the raw data is always recommended (IAEA, 2009b) because deleting values that correspond to inconsistent injections would improve both the precision and accuracy of LGR laser spectroscopes.

3.7 Repeatability

Seven samples were selected to assess the repeatability of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ measurements provided by the OA-ICOS instruments. These samples were analyzed three times (in three different vials) during the same run. Results for four representative samples with different isotopic composition are displayed in Figs. 7 and 8, with three repetitions for each instrument presented (in gray) along with the mean (in black). Error bars refer to the standard deviation computed for each measurement. IRMS $\delta^2\text{H}$ or $\delta^{18}\text{O}$ values and standard deviations are represented as horizontal solid and dashed lines, respectively. A visual inspection of the four instruments revealed inconsistent behaviour. The repeated measurements were very similar and within the instrumental precision in some cases (e.g., analyzer I in Fig. 7d and analyzer II in Fig. 8c) and appeared unsteady in other instances (e.g., analyzer I in Fig. 7b and analyzer IV in Fig. 8d). Particularly, repeated $\delta^2\text{H}$ measurements of sample 14 (Fig. 7, panel a) by spectroscopes I, II, and III fell within the analytical uncertainty of the IRMS, which resulted in differences between the lowest and the highest measurement equivalent to 0.71% , 0.57% , and 0.26‰ , respectively. These values were comparable or lower than the instrumental precision and revealed a satisfying repeatability of the instruments. In contrast, analyzer IV produced more unstable results with a greater difference between the lowest and highest $\delta^2\text{H}$ measurements (1.63‰). Table 6 presents the basic statistics of the maximum difference computed between repeated measurements of hydrogen and oxygen, respectively. Analyser II offered the best repeatability for both isotopic ratios, analyzers I and III yielded comparable results in terms of repeatability

Table 4. Basic descriptive statistics of standard deviation values of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ measurements.

LWIA	mean (‰)		std. dev. (‰)		min. (‰)		max. (‰)		25th percent. (‰)		75th percent. (‰)	
	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$
I	0.56	0.10	0.34	0.05	0.07	0.01	1.70	0.23	0.32	0.06	0.72	0.14
II	0.82	0.15	0.63	0.12	0.17	0.01	3.76	0.78	0.47	0.09	0.99	0.17
III	1.80	0.27	1.04	0.21	0.57	0.06	4.99	1.10	1.00	0.14	2.20	0.31
IV	1.01	0.12	0.54	0.07	0.19	0.03	2.84	0.46	0.54	0.08	1.34	0.14

**Fig. 7.** Repeatability plots of $\delta^2\text{H}$ measurements (A1 scheme). The mean value among the three repetitions is represented by the darker symbol. Solid and dashed horizontal lines represent the mass spectrometer $\delta^2\text{H}$ measurement and the standard deviation, respectively. Samples: (a) 14; (b) 25; (c) 23; (d) 24.

for $\delta^2\text{H}$ measurements, and analyzer IV exhibited the highest fluctuations in the repeated measurements. Analyzer IV behaved almost analogously to analyzer I for oxygen quantification, whereas analyzer III presented the most marked variations.

Overall, the capability to reproduce comparable results from the analysis of repeated samples was acceptable, with differences between the maximum and the minimum values which were generally within the range of the standard deviation yielded by the single measurements. This result agreed with previous studies of the LGR analyzers (IAEA, 2009b). Nevertheless, in some instances, the repeated measurements of the same sample were relatively different with marked unsteadiness and randomly distributed inconsistencies.

Table 5. Number of water molecules generated for four subsequent injections and the corresponding reportable delta values. In bold: the injection with an inconsistent number of water molecules.

injection number in the run	number of water molecules per cm^3 ($\times 10^{16}$)	$\delta^2\text{H}$ (‰)	$\delta^{18}\text{O}$ (‰)
255	3.55	-78.04	-10.97
256	3.54	-79.57	-11.15
257	3.13	-88.42	-12.10
258	3.72	-76.19	-10.43

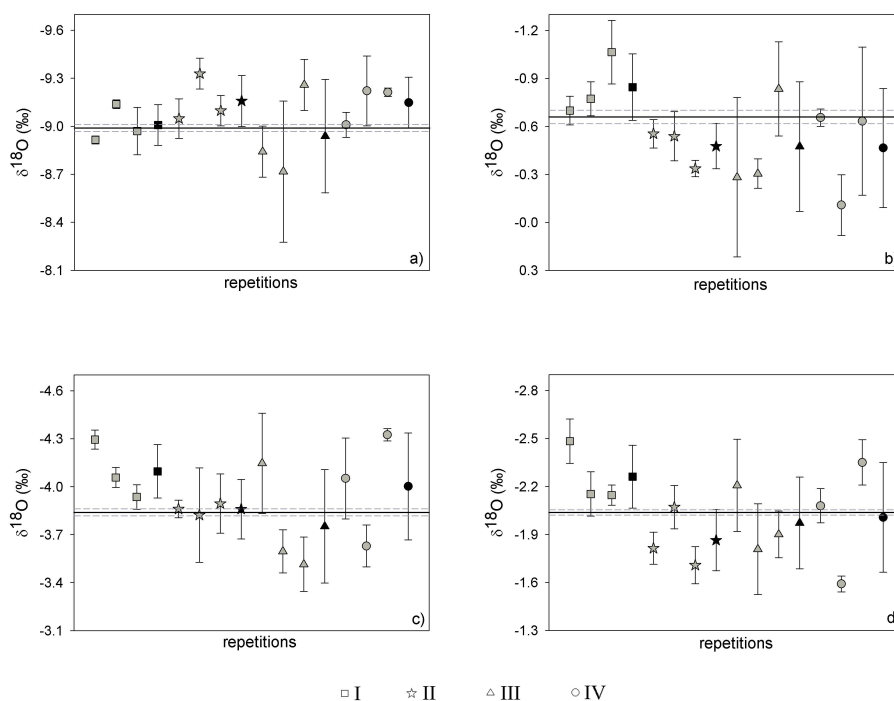


Fig. 8. Repeatability plots of $\delta^{18}\text{O}$ measurements (A1 scheme). The mean value among the three repetitions is represented by the darker symbol. Solid and dashed horizontal lines represent the mass spectrometer $\delta^{18}\text{O}$ measurement and the standard deviation, respectively. Samples: (a) 14; (b) 25; (c) 23; (d) 24.

Table 6. Basic descriptive statistics of the maximum difference between repeated $\delta^2\text{H}$ and $\delta^{18}\text{O}$ measurements.

LWIA	mean (‰)		std. dev. (‰)		min. (‰)		max. (‰)	
	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$
I	0.94	0.31	0.67	0.06	0.19	0.22	1.94	0.37
II	0.56	0.20	0.24	0.12	0.33	0.07	1.00	0.36
III	0.95	0.81	0.52	0.37	0.26	0.40	1.71	1.40
IV	2.21	0.44	1.01	0.23	0.23	0.17	3.24	0.76

Repeatability plots displayed in Figs. 7 and 8 also present the mean computed for three samples analysed over time (darker symbol). Averaging three repeated measurements may overcome the random deviations from the real isotopic ratios that are occasionally generated. The same statistical procedure was followed as in the analysis of $\delta^2\text{H}$ and $\delta^{18}\text{O}$ deviations (Sect. 2.4) to determine if this approach might lead to a significant improvement of results. The dataset investigated during this study was formed by the first, second, or third repetitions and by the mean value among the three repetitions for the seven samples. The distribution of each series was assessed by the Shapiro-Wilk, Kolmogorov-Smirnov, and Anderson-Darling normality tests at a significance level of 0.05. According to the type of distribution, a one-sample t-test (for normal distributions) or non parametric one-sample sign test (for the few non-normal distri-

butions) was performed to assess whether the deviations between the OA-ICOS and the IRMS measurements for the seven repeated samples were statistically significant. The t-test and sign test results are presented in Table 7 for hydrogen and oxygen measurements. In many cases, the laser spectrometers yielded accurate measurements for a single vial. However, averaging the values obtained by three identical samples almost always yielded reportable delta values that were not statistically different from IRMS ($\alpha:0.05$). For hydrogen, the deviation from the mass spectrometer values was significant in three of 12 cases, whereas the mean among three samples always produced more consistent results that were not significantly different from the reference value. For oxygen, the results deviated from the IRMS output in three of 12 cases, while the average almost always yielded (three times out of four) accurate results compared with mass

Table 7. Results of one-sample t-test and one-sample sign test for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ deviations of repeated samples. The analyses that produced negligible differences between laser spectroscopy and mass spectrometer (i.e., not significant difference with $\alpha:0.05$) are marked in bold.

LWIA	repetition n.	T-test for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ deviations				Sign test for $\delta^2\text{H}$ and $\delta^{18}\text{O}$ deviations			
		$\delta^2\text{H}$	$\delta^{18}\text{O}$	$\delta^2\text{H}$	$\delta^{18}\text{O}$	median	p -value	median	p -value
I	1	-1.41	0.21	-1.71	0.14	–	–	–	–
	2	0.25	0.81	-2.99	0.02	–	–	–	–
	3	-2.99	0.02	-1.20	0.28	–	–	–	–
	mean	-1.38	0.22	-3.12	0.02	–	–	–	–
II	1			0.80	0.46	0.30	0.12		
	2	1.10	0.31	-0.96	0.37	–	–	–	–
	3	3.91	0.01	1.33	0.23	–	–	–	–
	mean	2.22	0.07	0.42	0.69	–	–	–	–
III	1	0.12	0.91	0.40	0.70	–	–	–	–
	2	1.63	0.15	–	–	–	–	0.30	0.02
	3	0.02	0.98	-0.98	0.36	–	–	–	–
	mean	0.60	0.57	1.89	0.11	–	–	–	–
IV	1	1.69	0.14	-1.41	0.21	–	–	–	–
	2	3.77	0.01	1.47	0.19	–	–	–	–
	3	-0.87	0.42	-2.63	0.04	–	–	–	–
	mean	1.50	0.18	-0.52	0.62	–	–	–	–

spectrometry. According to these results, a higher degree of accuracy can be obtained with LGR analysers by averaging the measurements of three samples. In these cases, the increased analysis speed of spectroscopy IV (new version) allowed for more consistent results to be achieved by averaging values in a considerably shorter time frame than spectroscopes I, II, and III.

4 Conclusions

Because of their many advantages and great research potential, the use of off-axis integrated cavity output spectroscopy instruments is rapidly emerging among numerous institutions that deal with hydrological and natural resources research. Despite the number of previous investigations on the performance of such analyzers, no study to date has inspected the consistency of results obtained using different units of the same model. The present study focused on an inter-comparison of four liquid water isotope analyzers manufactured by Los Gatos Research Inc. (Mountain View, California, USA), which were versions 908-0008 and 908-0008-2000 of model DLT-100. This investigation aimed at assessing the performance of the four spectroscopes in terms of measurement reproducibility and repeatability in comparison with the performance of a traditional isotope-ratio mass spectrometer, on a wide range of isotopic ratios of natural water samples (ranging from -425‰ to -11‰ for $\delta^2\text{H}$ and

from -55‰ to -1‰ for $\delta^{18}\text{O}$). The laser units were operated running three different analytic schemes for the isotopic determination of water samples.

Scatterplots of laser-based versus IRMS-based measurements over the whole dataset demonstrated an excellent agreement between the two methods for both water isotopes and for all analyzers, which confirmed the overall good performance of OA-ICOS instruments, as had been indicated by previous studies. However, statistical analysis of deviations from the mass spectrometer measurements revealed a certain degree of variability in accuracy and precision among the four instruments and the two isotopic ratios. No bias or systematic deviations were evident for any particular machine and none was indicated as the overall best or worst performer. Nevertheless, one spectroscopy exhibited a marked positive deviation from zero for hydrogen measurements, whereas another analyzer consistently underestimated oxygen measurements. A third instrument lacked precision, especially for oxygen, compared with the other instruments. Interestingly, there was no causal relation between OA-ICOS - IRMS deviations and external factors, and the intrinsic variability of the analyzers was the only determined cause for such differences. Errors appeared to be randomly distributed within the same instrument and among the four machines. Therefore, the source of this variability remains unknown. The only evidence regarding a certain degree of error was related to the extremely light isotopic content of samples when $\delta^2\text{H}$ values were more negative than -300‰ , which resulted in a clear

underestimation of all instruments. Such a behaviour was much less marked for oxygen measurements and could be partly related to the substantial influence of memory effects when analyzing very light samples.

The use of different analytic schemes did not seem to significantly improve the isotopic measurements by laser spectroscopy. For all machines and both water isotopes, no scheme was able to consistently provide the most accurate and precise or worst measurements. Accuracy and precision seemed to be more related to the spectroscope than to the scheme. However, the analysis scheme first described by the International Atomic Energy Agency ranked among the best compared with IRMS.

Results also showed that the $1\text{-}\sigma$ precision ranged between $\pm 0.56\text{‰}$ and 1.80‰ for $\delta^2\text{H}$ and between $\pm 0.10\text{‰}$ and 0.27‰ for $\delta^{18}\text{O}$ measurements for the various instruments. Overall, these values were comparable or better than those reported by the manufacturer and in previous studies. One of the four analyzers yielded slightly more precise results for both isotopic ratios, but another instrument lacked precision unrelated to any evident factor.

Analyses conducted on a subset of samples revealed an acceptable capability of laser instruments to reproduce comparable results on repeated samples. The differences between maximum and minimum measurements generally fell within the range of the standard deviation of a single measurement. Averaging the delta values of three identical samples almost always led to a higher degree of accuracy and avoided potential random deviations. This approach was very time-consuming, and therefore might be applicable for the analysis of only a few samples and/or when more robust results are necessary.

In conclusion, OA-ICOS laser analyzers appeared to be cost-effective and not particularly difficult to operate compared with conventional mass spectrometry. Despite a certain degree of inter-machine variability and some randomly distributed errors, these instruments are a powerful approach for hydrological and environmental applications to determine hydrogen and oxygen isotopic compositions in water samples.

Acknowledgements. This work was supported by the STREP Project HYDRATE, Contract GOCE-037024, the Research Project GEO-RISKS (University of Padova, STPD08RWBY), the Czech Ministry of Education, Czech Republic (CEZ MSM 6840770002), and the National Research Fund of Luxembourg (TR-PHD BFR07-047-Investigating scale dependency of runoff generation). We would like to express special thanks to Jeff McDonnell (Oregon State University, Corvallis, USA) for useful feedback at the beginning of this study, Tyler Coplen for providing the USGS RSIL W-64444 sample of Antarctic water (sample 30), François Barnich (Centre de Recherche Public – Gabriel Lippmann, Luxembourg) for continuous support during sample and data analysis, Luigi Dallai (National Council Research, Pisa, Italy) for analysis of the first calibration samples of the Padova spectroscope, and Markus Weiler (University of Freiburg, Germany) for interesting discus-

sions about isotopic analyses. We also thank Brent D. Newman and Luis Araguás-Araguás (IAEA, Vienna) and an anonymous reviewer for their valuable comments that greatly improved the final version of the manuscript.

Edited by: S. Uhlenbrook

References

- Aggarwal, P. K., Ahmad, T., Groening, M., Gupta, M., Owano, T., Baer, D.: Laser spectroscopic analysis of stable isotopes in natural waters: a low-cost, robust technique for the use of environmental isotopes in hydrological and climate studies, *Eos Trans. AGU*, 87(52), Fall Meeting Supplement, Abstract H51D-504, 2006.
- Baer, D. S., Paul, J. B., Gupta, M., and O'Keefe, A.: Sensitive absorption measurements in the near-infrared region using off-axis integrated-cavity-output spectroscopy, *Appl. Phys.*, B75, 261–265, 2002.
- Berman, E. S. F., Gupta, M., Gabrielli, C., Garland, T., and McDonnell, J. J.: High-frequency field deployable isotope analyzer for hydrological applications, *Water Resour. Res.*, 45, W10201, doi:10.1029/2009WR008265, 2009.
- Brown, V. A., McDonnell, J. J., Douglas, A. B., and Kendall, C.: The role of event water, a rapid shallow flow component, and catchment size in summer storm flow, *J. Hydrol.*, 217, 171–190, 1999.
- Epstein, S. and Mayeda, T. K.: Variations of ^{18}O of waters from natural sources, *Geochim. Cosmochim. Ac.*, 4, 213–224, 1953.
- Gonfiantini, R.: Standards for stable isotope measurements in natural compounds, *Nature*, 271, 534–536, 1978.
- Horita, J. and Kendall, C.: Stable isotope analysis of water and aqueous solutions by conventional dual-inlet mass spectrometry. In: P.A. de Groot (Ed.), *Handbook of Stable Isotope Analytical Techniques*, Elsevier, Amsterdam, 1–37, 2004. 2004.
- Horita, J., Ueda, A., Mizukami, K., and Takatori, I.: Automatic δD and $\delta^{18}\text{O}$ analyses of multi-water samples using H_2 - and CO_2 -water equilibration methods with a common equilibration set-up, *Appl. Radiat. Isotopes*, 40(9), 801–805, 1989.
- Hut, G.: Consultants' group meeting on stable isotope reference samples for geochemical and hydrological investigations. Scientific Meeting 16–18 Sept. 1985, IAEA, Vienna Internal Report, p. 42, 1987.
- Huth, A. K., Leydecker, A., Sickman, J. O., and Bales, R. C.: A two-component hydrograph separation for three high-elevation catchments in the Sierra Nevada, California, *Hydrol. Process.*, 18, 1721–1733, 2004.
- IAEA: Reference Sheet for VSMOW2 and SLAP2 international measurement standards. Issued 2009-02-13, International Atomic Energy Agency, Vienna, 5 p., <http://nucleus.iaea.org/rpst/Documents/VSMOW2.SLAP2.pdf>, 2009a.
- IAEA: Laser Spectroscopy Analysis of Liquid Water Samples for Stable Hydrogen and Oxygen Isotopes. Performance Testing and Procedures for Installing and Operating the LGR DT-100 Liquid Water Isotope Analyzer, International Atomic Energy Agency, Vienna, 2009, ISSN 1018-5518, 2009b.
- Koeniger P., Hubbart A. J., Link T., Marshall J. D., Isotopic variation of snow cover and streamflow in response to changes in

- canopy structure in a snow-dominated mountain catchment, *Hydrol. Process.*, 22, 557–566, 2008.
- La Bolle, E. M., Fogg, G. E., Eweis, J. B., Gravner, J., and Leaist, D. G.: Isotopic fractionation by diffusion in groundwater, *Water Resour. Res.*, 44, W07405, doi:10.1029/2006WR005264, 2008.
- Lee, K., Kim, J., Lee, D., Kim, Y., and Lee, D.: Analysis of water movement through an unsaturated soil zone in Jeju Island, Korea using stable oxygen and hydrogen isotopes, *J. Hydrol.*, 345, 199–211, 2007.
- Lis, G., Wassenaar, L. I., and Hendry M. J.: High precision laser spectroscopy D/H and $^{18}\text{O}/^{16}\text{O}$ measurements of microliter natural water samples, *Anal. Chem.*, 80, 287–293, 2008.
- Los Gatos Research, Inc.: Liquid-Water Isotope Analyser, Automated Injection, 2008.
- Los Gatos Research, Inc.: Liquid-Water Isotope Analyser, Version 2. Automated Injection, 2010.
- Lyon, S. W., Desilets, S. L. E., and Troch, P. A.: Characterizing the response of a catchment to an extreme rainfall event using hydrometric and isotopic data, *Water Resour. Res.*, 44, W06413, doi:10.1029/2007WR006259, 2008.
- Lyon, S. W., Desilets, S. L. E., and Troch, P. A.: A tale of two isotopes: differences in hydrograph separation for a runoff event when using δD versus $\delta^{18}\text{O}$, *Hydrol. Process.*, 23, 2095–2101, 2009.
- McGuire, K. J. and McDonnell, J. J.: A review and evaluation of catchment transit time modeling, *J. Hydrol.*, 330, 543–563, 2006.
- Paul, J. B., Lapson, L., and Anderson, J. G.: Ultrasensitive absorption spectroscopy with a high-finesse optical cavity and off-axis alignment, *Appl. Optics*, 40(27), 4904–4910, 2001.
- Ricci, R. W., Ditzler, M., and Nestor, L. P.: Discovering the Beer-Lambert law, *J. Chem. Educ.*, 71(11), 983–985, 1994.
- Rodgers, P., Soulsby, C., and Waldron, S.: Stable isotope tracers as diagnostic tools in upscaling flow path understanding and residence time estimates in a mountainous mesoscale catchment, *Hydrol. Process.*, 19, 2291–2307, 2005a.
- Rodgers, P., Soulsby, C., Waldron, S., and Tetzlaff, D.: Using stable isotope tracers to assess hydrological flow paths, residence times and landscape influences in a nested mesoscale catchment, *Hydrol. Earth Syst. Sci.*, 9, 139–155, doi:10.5194/hess-9-139-2005, 2005.
- Roether, W.: Water- CO_2 exchange set-up for the routine oxygen-18 assay of natural waters, *International Journal of Applied Radiation and Isotopes*, 21, 379–387, 1970.
- Rolston, J. H., den Hartog, J., and Butler, J. P.: The deuterium isotope separation factor between hydrogen and liquid water, *J. Phys. Chem.-US*, 80, 1064–1067, 1976.
- Sayres, D. S., Moyer, E. J., Hanisco, T. F., St. Clair, J. M., Keutsch, F. N., O'Brien, A., Allen, N. T., Lapson, L., Demusz, J. N., Rivero, M., Martin, T., Greenberg, M., Tuozzolo, C., Engel, G. S., Kroll, J. H., Paul, J. B., and Anderson, J. G.: A new cavity based absorption instrument for detection of water isotopologues in the upper troposphere and lower stratosphere, *Rev. Sci. Instrum.*, 80, 044102, doi:10.1063/1.3117349, 2009.
- Singleton, G. L., Coplen, T. B., Qi, H., and Lorenz, J. M.: Laser-based stable hydrogen and oxygen analyses: how reliable can measurements results be? EGU General assembly 2009, *Geophysical Research Abstracts*, 11, EGU2009-3290, 2009.
- Stewart, M. K., Morgenstern, U., and McDonnell, J. J.: Truncation of stream residence time: how the use of stable isotopes has skewed our concept of stream water age and origin, *Hydrol. Process.*, 24, 1646–1659, doi:10.1002/hyp.7576, 2010.
- Taylor, S., Feng, X., Williams, M., and McNamara, J.: How isotopic fractionation of snowmelt affects hydrograph separation, *Hydrol. Process.*, 16, 3683–3690, 2002.
- Tetzlaff, D., Soulsby, C., Waldron, S., Malcolm, I. A., Bacon, P. J., Dunn, S. M., Lilly, A., and Youngson, A. F.: Conceptualization of runoff processes using a geographical information system and tracers in a nested mesoscale catchment, *Hydrol. Process.*, 21, 1289–1307, 2007.
- Uhlenbrook, S. and Hoeg, S.: Quantifying uncertainties in tracer-based hydrograph separations: a case study for two-, three- and five-component hydrograph separations in a mountainous catchment, *Hydrol. Process.*, 17, 431–453, 2003.
- Wang, L., Caylor, K. K., and Dragoni, D.: On the calibration of continuous, high-precision $\delta^{18}\text{O}$ and $\delta^2\text{H}$ measurements using an off-axis integrated cavity output spectrometer, *Rapid Commun. Mass. Sp.*, 23, 530–536, 2009.
- Wassenaar, L. I., Hendry, M. J., Chosten, V. L., and Lis, G. P.: High resolution pore water $\delta^2\text{H}$ and $\delta^{18}\text{O}$ measurements by $\text{H}_2\text{O}(\text{liquid})\text{-H}_2\text{O}(\text{vapor})$ equilibration laser spectroscopy, *Envir. Sci. Tech.*, 42, 9262–9267, 2008.
- Weiler, M., McGlynn, B. L., McGuire K. J., and McDonnell, J. J.: How does rainfall become runoff? A combined tracer and runoff transfer function approach, *Water Resour. Res.*, 39, 1315, 4–1, 4–13, doi:10.1029/2003WR002331, 2003.
- Zhou, S., Nakawo, M., Hashimoto, S., and Sakai, A.: The effect of refreezing on the isotopic composition of melting snowpack, *Hydrol. Process.*, 22, 873–882, 2008.