

Author's reply to the referees comments to manuscript AMT-2017-120 - Anonymous referee 3

The original referee's comments are written in black and the author's reply and changes to the manuscript are colored in blue/green respectively. References to page and line numbers as well as figures refer to the original manuscript, but references to sections refer to the corrected manuscript. In cases where we insert figures, tables and equations into this document, they are referenced with R1, R2, R3 ...

Interesting paper but needs some major revisions. Please find below some listed points that should be changed or at least answered.

Author's response: We thank the anonymous referee for the feedback and the suggestions to our manuscript, below we answer the referee's comments.

1a) Page 2 lines 13ff: text passage about IRMS: Pls cite Schnyder et al. there (citation below)

1b) in the same text passage: I think "sample preparation effort and cost" might be a minus for IRMS techniques. But here the main disadvantages should be mentioned like (storage) problems with vials (see Gemery et al., 1996 and Knohl et al., 2004) and the advantage of quasi-continuous measurements relative to the "discontinuous" measurement by IRMS.

Author's response/changes to the manuscript: Thanks for these suggestions, we added this information to the respective paragraph: "IRMS has been widely used for isotope studies in environmental sciences, but shows limited applicability for *in situ* measurements (Griffis, 2013), but see also the field applicable continuous flow IRMS described by (Schnyder, 2004). Disadvantages of flask-sampling based IRMS techniques include high sample preparation effort and costs (Griffis, 2013), low temporal resolution and discontinuous measurements. Additionally, there are potential problems during sample storage and transport, see (Knohl et al 2004) for minimizing such storage effects in case of ^{13}C . For ^{18}O storage effects can be related to oxygen exchange between water and CO_2 (Gemery 1996, Tuzson 2008)."

2) Page 2 lines 22ff. text passage about different spectrometer types: should be shortened as this manuscript is not a review on optical methods for measurement of isotope ratios

Author's response/changes to the manuscript: We shortened the introduction in especially page 2 lines 22ff.

3) Page 3 lines 25ff: "to characterize the Delta Ray IRIS and its performance under field conditions": I think measurement of the "internal cell turnover" and "Allan deviation" is not sufficient to fulfill this topic here. The reference gas box from the Delta Ray is said to offer possibilities to adjust CO_2 conc of the "reference" gas to the measured $[\text{CO}_2]$ to cancel out a possible concentration effects on the measured δ -values. The authors need to go more in detail here by showing data (!) from multiple CO_2 -in-air-standards with different $[\text{CO}_2]$ and different $\delta^{13}\text{C}$ -and $\delta^{18}\text{O}$ values measured with IRMS (preferred) in comparison to measurement with Delta Ray or a comparison with different optical measurement devices (more problematic). I suppose you have measured the data, so show them here please.

Author's response: We show the measured concentration dependency as well as a comparison of multiple CO_2 -in-air standards with IRMS measurements (δ values) and measurements with a Picarro (concentration) in figures R1 R2 and R3.

We added the description of the additional measurements to the manuscript:

2.6 Instrument characterization measurements

We carried out additional measurement in the field and in the lab to quantify precision, evaluate the calibration strategy and quantify the instrument's response time and repeatability. These measurements involved changes in the analyzers plumbing. For all measurements that required connecting different gas tanks to the analyzer, they were either connected directly to the analyzer's

internal ports ('CRef1' and 'CRef2') or the plumbing was equivalent to the plumbing of the target gas (Fig.1).

- 1) Lab measurements to quantify precision and evaluate the calibration strategy
 - We measured the Allan deviation by connecting pressurized air at atmospheric δ values to the analyzer and took measurements at the analyzer's maximum data acquisition rate of 1 Hz for two hours.
 - We diluted pure CO₂ with synthetic air over a CO₂ concentration range of 200 to 1500 ppm to measure the concentration dependency of the measured (raw) δ values. This dilution experiment was carried out for three different tanks with pure CO₂ at different δ values. Each gas tank was measured twice. (Used gas tanks: "ambient", "bio1" and "bio2", c.f. Table 3.)
 - We measured the concentration c and the isotopic compositions $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ of gases with concentrations ranging from (350 to 450 ppm) and isotopic compositions ranging from -37 to -9.7 ‰ for $\delta^{13}\text{C}$ and from -35 to -5 ‰ for $\delta^{18}\text{O}$. Each of these measurements was performed three times. (Used gas tanks: "ambient", "bio1", "bio2", "PA-tank", SACO₂ -350, SACO₂ -450, SACO₂ -500, c.f. Table 3.)
 - We performed measurements of two pure CO₂ gas tanks at different δ values (diluted to different concentrations between 200 and 3000 ppm) as well as measurements of two gas tanks at different CO₂ concentrations (350 and 500 ppm). These measurements were repeated every six hours for a period of one week. (Used gas tanks: "ambient", "bio", ('SA-CO₂-350' and 'SA-CO₂-500', c.f. table 3.)"
- 2) Field measurements to quantify the setup's response time and repeatability
 - The response time of the tubing and the analyzer was measured by using the automatic switching unit (Figure 1) to switch from ambient air (height 1) to the target standard. We superimposed the measurements of four switching events to observe the adjacent turnover processes.
 - The analyzer's repeatability under field conditions was quantified by the half hourly target measurements described in Sect. 2.5.

We removed the chapter "Accuracy" and replaced it by the following:

3.1.2 Evaluation of the calibration strategy

The instrument's internal calibration strategy (described in section 2.7.1) is based on:

- A nonlinear relationship between raw δ values and concentrations (Figure R1).
- A linear relationship between calibrated δ value (measured with IRMS) and the concentration-corrected δ value - $\delta_{c\text{-corrected}}$ in Equation R1 (Fig. R2, left panel).
- A linear relationship between measured (raw) and real concentrations (Figure R2, middle and right panel).
- The repeatability of the calibration curves – for δ values modulo the Offset correction, that is applied by the instrument's internal 'Referencing' (Figure R3 and Table R1).

Raw δ values show a nonlinear dependency from raw concentrations (Fig. R1). This nonlinear relationship deviates from the concentration-dependency correction applied by the instrument ($\delta_{c\text{-corrected}}$ in equation R1). In Fig. R1, this function is shown for the used gas tank 'ambient' after an Offset correction at a concentration at 400 ppm, which is similar to the instrument's internal 'referencing'. Thus, the deviations of the measured δ values from the concentration-dependency correction (top panel of Fig. R1) give an estimate of the uncertainty of measurements that is related to the deviation from the reference concentration. For referencing at 400 ppm, these deviations were below 0.2 ‰ for ¹³C and 0.4 ‰ for ¹⁸O.

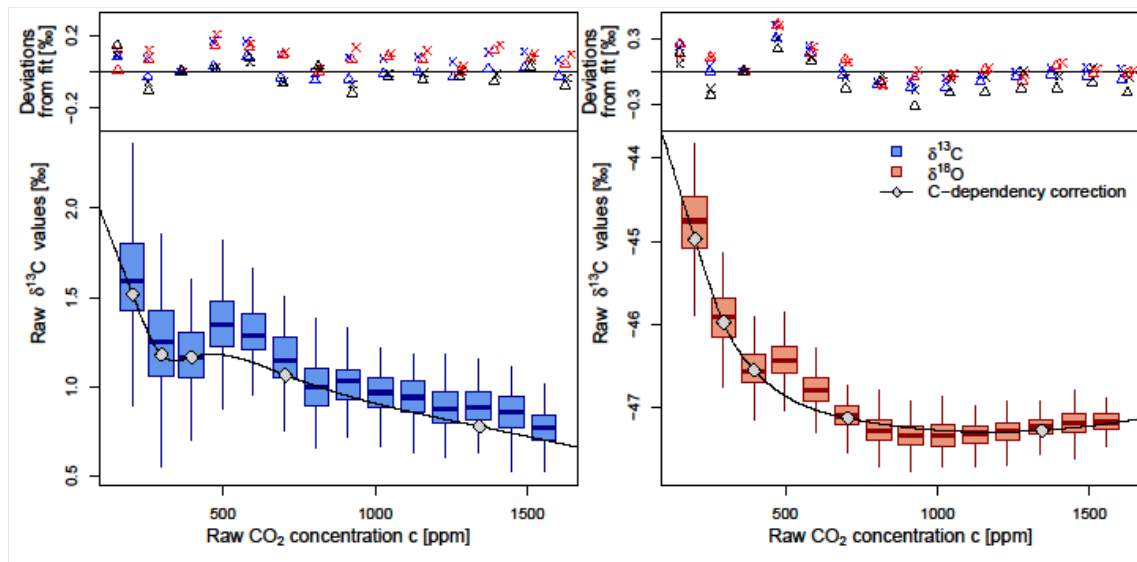


Figure R1 Box whiskers plots showing the nonlinear concentration-dependency of raw δ values for ^{13}C and ^{18}O respectively, here as an example for the CO_2 tank 'ambient'. This measured c -dependency is compared to the respective concentration-dependency correction (black line, with grey symbols marking the data points used during the respective calibration measurement). The c -dependency correction is Offset-corrected to match the raw δ values at 400ppm and the mean deviation from the fit is shown in the top panel for two measurements (different symbols) with three different gas tanks ('ambient' in blue, 'bio' in black and 'bio2' in red).

The measured linear relationships for concentration and δ scale calibration (Fig. R2) have R^2 values of above 0.9999 for concentration, above 0.999 for $\delta^{13}\text{C}$, and above 0.998 for $\delta^{18}\text{O}$. The linearity and potential accuracy, as defined by (Tuzson et. al., 2008) can be quantified as the 1σ standard deviation from the linear fits. The so defined potential accuracy of the instrument internal calibration is 0.45 ppm for CO_2 concentration; 0.24 ‰ for $\delta^{13}\text{C}$ and 0.3 ‰ for $\delta^{18}\text{O}$. For both δ values, this is comparable to the uncertainty related to the nonlinear concentration calibration that varies with δ and c as discussed above.

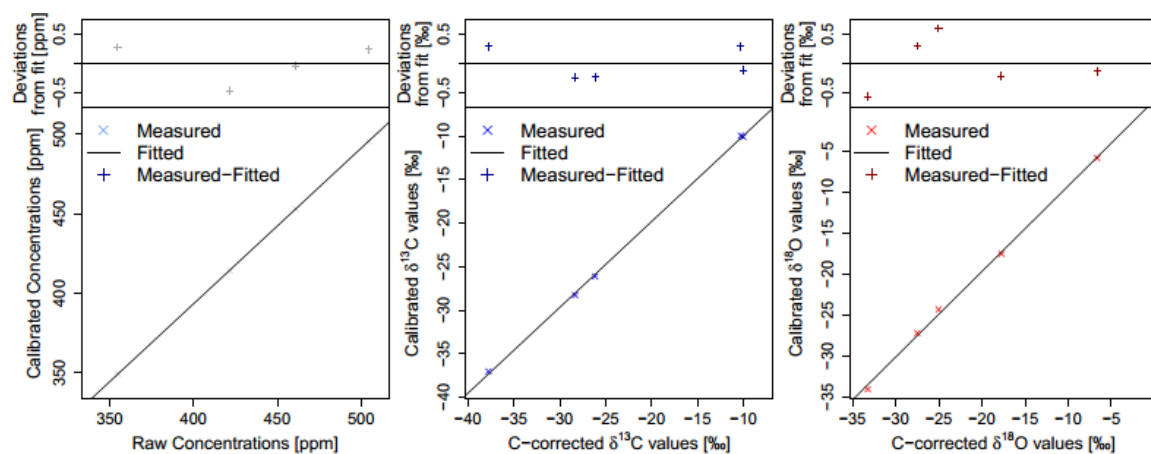


Figure R2 Linear calibrations for concentration (left panel), $\delta^{13}\text{C}$ (middle panel) and $\delta^{18}\text{O}$ (right panel).

The repeatability of the calibration curves is discussed here based on measurements of the nonlinear concentration dependency (Figure R1), and repeated measurements of gas tanks with two different c

and δ values to evaluate temporal changes in the respective linear relationships (Figure R2). These measurements were taken every six hours for a period of nine days. The standard deviation of the different measurement is below 0.2 ppm for concentrations and below 0.05 and 0.1 ‰ for ^{13}C and ^{18}O respectively. Thus the uncertainty related to the repeatability of the linear calibrations is smaller than the potential accuracy discussed above. For δ values, these values are comparable to the repeatability reported by several authors measured with other laser spectrometers (e.g. Sturm et al 2011; 2013; Vogel et al 2013). For concentrations on the other hand, Sturm et al 2013 reported a much smaller value of 0.03 ppm, based on more frequent calibration. In our setup, the concentration calibration is only performed once after the instrument is restarted, thus there might be a potential for better repeatability in concentration measurements by more frequent concentration calibration. For δ values, the repeatability that is related to deviations from the reference concentration depends on concentration (Table R1). Repeated measurements of these deviations have standard deviations of below 0.15 ‰ for concentrations between 200 and 1600 ppm.

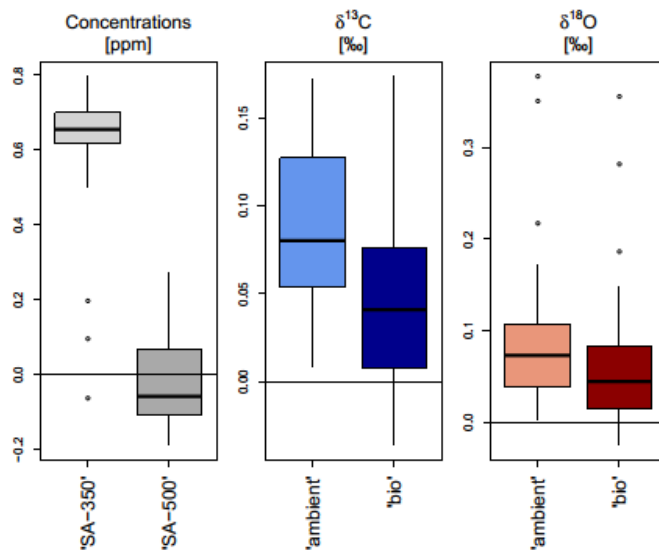


Figure R3 Box whiskers plots for the deviations of calibrated concentrations and δ values from laboratory measurements (at MPI in Jena) for repeated measurements of different calibration tanks (c.f. Table 3 for c and δ values of the gas tanks) over a period of 9 days ($N=36$). Delta values were measured at 400 ppm and ‘referencing’ was done app. Every 30 minutes at 380ppm to simulate conditions during a measurement campaign.

	tank ,ambient‘		tank ,bio‘	
Concentrations	$\sigma (\delta^{13}\text{C}- \delta^{13}\text{C}_{\text{tank}})$	$\sigma (\delta^{18}\text{O}- \delta^{13}\text{C}_{\text{tank}})$	$\sigma (\delta^{13}\text{C}- \delta^{13}\text{C}_{\text{tank}})$	$\sigma (\delta^{18}\text{O}- \delta^{18}\text{O}_{\text{tank}})$
202	0,07	0,14	0,09	0,13
396	0,04	0,05	0,08	0,08
600	0,09	0,08	0,12	0,12
807	0,08	0,08	0,11	0,11
1018	0,10	0,08	0,13	0,11
1232	0,12	0,09	0,13	0,11
1450	0,14	0,11	0,15	0,12
1664	0,14	0,11	0,14	0,12
3145	0,17	0,15	0,17	0,15

Table R1 Standard deviations σ of the differences between the calibrated δ values and the known values of used tanks ‘ambient’ and ‘bio’ over a large concentration range.

4) Please give more info (citation if available) on the kind of measurements performed at the MPI in Jena (isotopes and concentration).

Author's response/changes to the manuscript: We added this information to the manuscript:

"All CO₂ containing gas tanks that were used for calibration as well as the target gas tank were measured high precisely for their CO₂ concentration and isotopic composition in ¹³C and ¹⁸O at the Max Planck Institute for Biogeochemistry in Jena. There, the CO₂ concentrations were measured with a Picarro CRDS G1301 and the isotopic composition was measured with IRMS linked to VPDB (VPDB-CO₂) by using the multi point scale anchor JRA-S06 (Wendeberg et al 2013)."

5) The link to VPDB was done with the gas tank measured in Jena? Please extend the info on how this is done. Fig. 3 describes your quality control standard? Is there a way to compare measured values (+stdev.) with a target value (+stdev.)?

Author's response:

Concerning the link to VPDB: We added a chapter that describes the calibration using the tanks that were measured in Jena.

Concerning Fig 3: We used the deviations between measured and target value with the respective uncertainties to calculate accuracy in chapter 3.1.2 and Fig 3 of the original manuscript. We changed this and used measurements with more than one tank to quantify 'potential accuracy' of the instrument.

2.6 Instrument internal calibration

The Delta Ray analyzer is equipped with three different internal calibration routines (Thermo Fisher Scientific, 2014). We performed these routines at the field site (in situ) each time the analyzer had to be restarted e.g. after power supply failures, instrument issues or when we manually turned off the analyzer for other reasons. All three instrument internal calibration procedures were usually done one day after restarting the analyzer, thus the instrument was in thermal equilibrium during calibration. The three different instrument internal calibration procedures are described below:

- **Correction of concentration dependency (called 'linearity calibration' in the instrument's documentation and operational software)**

This calibration routine evaluates the concentration dependency of δ value measurements (Thermo Fisher Scientific, 2014). Mathematically, an experimentally derived correction factor $f_{\text{correct}}(C_{\text{raw}})$ is multiplied with the raw isotopic ratio R (information from the manufacturer, Thermo Fisher Scientific)

$$R_{\text{c-corrected}} = f_{\text{correct}}(C_{\text{raw}}) \times R_{\text{raw}} \quad (\text{Equation R1})$$

This factor as a function of concentration is determined via a natural spline fit of measurements of a gas tank with constant δ value at different concentrations (information from the manufacturer, Thermo Fisher Scientific). This is implemented by mixing pure CO₂ with CO₂-free air, yielding concentrations between 200 to 3500 ppm. In our setup we used the pure CO₂ with near to ambient δ values (tank 'ambient CO₂', c.f. Table 3) and synthetic air for this calibration.

[...]

The instrument's internal calibration procedure is based on the measurement of these calibration curves after the instrument is started in combination with repeated measurements of a known gas, so called 'referencing' (see below). As the different calibrations are only performed once after the

instrument is restarted, the accuracy and repeatability of measurements is further based on the assumption that, these relationships remain sufficiently constant, and temporal changes are corrected by 'referencing'.

- **Referencing**

This procedure applies an offset correction of the calibrated δ values using a gas with known δ values that is measured at a freely selectable concentration in regular intervals (information from the manufacturer, Thermo Fisher Scientific). In our experimental setup, referencing is carried out every 30 minutes for 80 s after the tubes have been purged for 60 s using the pure CO₂ standard ('ambient CO₂', c.f. Table 3) diluted with synthetic air. We chose the reference concentration to be the same as in the highest inlet in the adjacent cycle, because most of the measurement inlets had concentrations close to those at the highest inlet and the temporal variability of the measured concentrations generally decreased with height. Thus, we performed the 'Referencing' as close as possible to as many height measurements as possible by these settings."

Thus, the calibration procedure for δ values can be expressed with the following formula with the correction factor $f_{\text{correct}}(C_{\text{raw}})$ as determined from the concentration dependency correction, and the slope $m_{\delta\text{scale}}$ derived from the δ scale calibration (information from the manufacturer, Thermo Fisher Scientific).

$$\delta_{\text{calibrated}}(R_{\text{raw}}; C_{\text{raw}}; t) = m_{\delta\text{scale}} \times \underbrace{\left(\frac{f_{\text{correct}}(C_{\text{raw}}) R_{\text{raw}}}{R_{\text{std}}} - 1 \right)}_{\delta_{\text{e-corrected}}} + \delta_{\text{Offset}}(t) \quad (\text{Equation R2})$$

6) Page 3 line 26 "b)" please add one or two sentences why $R_{\text{eco}}^{13\text{C}}$ and $R_{\text{eco}}^{18\text{O}}$ is interesting.

Author's response/changes to the manuscript: We added one more sentence to the first paragraph in the introduction. "The ¹³C composition of ecosystem respiration $R_{\text{eco}}^{13\text{C}}$ on the one hand, has been used to assess the time lag between assimilation and respiration (e.g. Ekblad and Höglberg, 2001; Bowling et al., 2002; Knohl et al., 2005) and to evaluate biosphere models on global scale (Ballantyne et al, 2011). The ¹⁸O composition of ecosystem CO₂ exchange $R_{\text{eco}}^{18\text{O}}$ on the other hand is particularly interesting to study the coupled CO₂ and water cycle (see e.g. Yakir and Wang, 1996)."

7) Page 11 line 21 "lighter" here means only ¹³C-depleted or also ¹⁸O-depleted ? Please specify (also in whole manuscript)

Author's response/changes to the manuscript: We specified this terminology throughout the manuscript.

8) Page 13 line 26: more "enriched" in what? Please check that also in whole manuscript, depleted in ¹³C, enriched in ¹⁸O (page 14 line 21...)

Author's response/changes to the manuscript: We specified this terminology throughout the manuscript.

9) I'm not totally happy to read a manuscript with 2 hypotheses where one hypothesis can be discarded but the 2nd one cannot be proven. The authors should find a way around this, at least the additional measurements for finally testing should be mentioned and discussed here.

Author's response/changes to the manuscript: We added a paragraph that describes which measurements would be needed to support this hypothesis. These measurements are however very

laborious, carry high uncertainty by themselves, and beyond the scope of this study. “To test this hypothesis, we would need to measure the amount and the isotopic composition of autotrophic respiration, total soil respiration and ecosystem respiration (e.g. by a trenching experiment) at our field site with an appropriate time resolution to capture the day-to-day variability during the field campaign. Lab measurements using incubations could also give an idea of the isotopic composition of autotrophic and total soil respiration, but would not fully reflect field site conditions. “

10) the unit “‰” is not conform to the SI unit system, what about using “mUr”? It might be more an editorial decision

Author’s response: As ‰ is so commonly used and also the literature we are citing uses ‰, we think it might be most convenient for the reader if we also use ‰, even if it is not a SI unit. We are however happy to follow the editor’s suggestion.

We added the suggested references to the manuscript:

Gemery et al. (1996): Oxygen isotope exchange between carbon dioxide and water following atmospheric sampling using glass flasks. J Geophys Res 101, D9, 14514-14420.

Knohl et al. (2004): Kel-FTM discs improve storage time of canopy air samples in 10-mL vials for CO₂-d¹³C analysis. Rapid Comm Mass Spectrom. 18, 1663-1665.

Schnyder et al. (2004): Mobile, outdoor continuous-flow isotope-ratio mass spectrometer system for automated high-frequency ¹³C- and ¹⁸O-CO₂ analysis for Keeling plot applications. Rapid Comm Mass Spectrom. 18, 3068-3074.