

Reply to Interactive comment of Referee #1 on “LISA: a lightweight stratospheric air sampler” by Joram J. D. Hooghiem et al.

We thank the two reviewers for their helpful and detailed comments on our manuscript. We have addressed the major comments from both reviewers regarding our derivation of the bias estimation in isotopic composition measurements and revised the introduction according to the comments of both reviewers. Below, a detailed point-to-point reply and a revised version of our manuscript with track changes are given.

The paper by Hooghiem presents a new technique for sampling of stratospheric air which may be well suited to add to the available techniques of cryogenic whole air sampling and AirCore. The paper is well written, although the introduction is a bit like a collection of information on sampling techniques and the use of stratospheric trace gas measurements but lacks a clear argumentation line. The subject is very much in line with the scope of AMT. I have a range of minor suggestions/questions and one major observation. The major observation explained below should be clarified before publication.

Major comments:

Major issues My main issue is the discussion of the differences between AirCore and LISA. First, it would be extremely valuable to see the paper by Chen et al., describing the data evaluation of the AirCore system used here. More importantly, I have doubts about the way that the fractional contribution of a contamination source (f_c) is derived in section 6.4. In my view, calculation of f_c from observed CO_2 [CO_2] and deviations between this observation and the expected stratospheric value ($[\text{CO}_2]_s$) should be calculated as follows (values in parenthesis are mixing ratios):

*$$[\text{CO}_2] = f_s * [\text{CO}_2]_s + f_c * [\text{CO}_2]_c = (1-f_c) * [\text{CO}_2]_s + f_c * [\text{CO}_2]_c$$
 With the subscripts as defined in the paper. Solving this for f_c yields: $f_c = ([\text{CO}_2] - [\text{CO}_2]_s) / ([\text{CO}_2]_c - [\text{CO}_2]_s)$ If I assume that $([\text{CO}_2] - [\text{CO}_2]_s)$ is the difference of 0.84 ppm CO_2 between observed CO_2 in LISA and in AirCore, I need to make an assumption on $([\text{CO}_2]_c - [\text{CO}_2]_s)$ to calculate f_c . If we assume that $[\text{CO}_2]_s$ is stratospheric CO_2 at about 390 ppm (actually 395 is more realistic), and that the contamination is from tropospheric CO_2 with a mixing ratio of around 405 ppm, then I calculate f_c to be $0.84/(405-395) = 0.084$, and not 0.0021 as derived in Table 6.*

We thank the reviewer for pointing out the calculation error. We mistakenly assumed that the contamination results from one-way diffusion of ambient CO_2 into the bag. If the observed bias would be entirely caused by diffused contamination air, the fraction of the contamination air is indeed about 8 percent. However, the observed bias is unlikely caused by diffusion alone based on our laboratory experiment results, where we found much smaller biases (~ 0.1 ppm for CO_2 and ~ 2 ppb for CH_4), as shown in Figure 2.

We have further thoughts on the diffusion process. According to the Fick's law, the diffusion depends on the diffusivity and the concentration gradient, which may be different for different species, e.g. the concentration gradient for O_2 and N_2 across the bag's material is negligible, and the concentration gradients for CO_2 and CH_4 depend on their ambient concentrations. We have updated the manuscript with the following analysis of the problem.

“The stratospheric air samples can be used for analysis of isotopic composition measurements of trace gases. Here we take CO_2 and CH_4 as an example to estimate the uncertainties of isotopic composition measurements due to the storage bias (see Table 3) or the AirCore-LISA

bias (see Table 5), and the estimated isotopic signatures associated with the assumed contamination source.

For any species, the measured number of molecules n_m , is the sum of the number of molecules from the original source, n_s , and the contamination n_c that entered the sampling bag through diffusion:

$$n_m = n_s + n_c \quad (6)$$

Diffusion is governed by Fick's law:

$$J = D \frac{\partial C}{\partial x} \quad (7)$$

Where J is the diffusion flux, D the diffusivity and $\frac{\partial C}{\partial x}$ the concentration gradient of diffusing species, e.g. CO_2 . Each layer of the MLF bag has its own specific diffusivity and sorption characteristics, and for simplicity here we assume that it can be modelled with a single diffusivity constant that is only species-dependent. The process of sorption of gases into the solid material and the interaction governing that process is also ignored for simplicity. Then the total diffusion out of the bag which is of stratospheric origin, denoted by y_{out} . The total amount diffusing into the bag, denoted by y_{in} , is of tropospheric origin. Assuming that these two fluxes can be modelled with a law of mass action:

$$y_{\text{in}} = DC_c \quad (8)$$

and

$$y_{\text{out}} = DC_s \quad (9)$$

Where C_c is the concentration of ambient air that contaminates the stratospheric sample in the bag denoted by C_s . The bias that we measure is as follows:

$$\text{bias} = y_{\text{in}} - y_{\text{out}} \quad (10)$$

Assuming that y_{out} is purely stratospheric and y_{in} is purely tropospheric, and that c_{out} and c_{in} are mole fractions outside and inside the bags and are assumed to be constant. We can estimate the fraction of tropospheric contamination in our sample. The measurement (C_m) is the sum of the original stratospheric sample (C_s), minus the sample that is leaving the bag and the tropospheric sample entering the bag:

$$C_m = C_s - y_{\text{out}} + y_{\text{in}} \quad (11)$$

And the contaminating fraction is then f_c is

$$f_c = \frac{y_{\text{in}}}{C_m} = \frac{C_c \cdot \text{bias}}{C_m(C_c - C_s)} \quad (12)$$

Where the right-hand side is obtained by using Eq. (8) to (10). The sample fraction, f_s , is simply $f_s = 1 - f_c$ (13)

The isotope composition after the mixing of the tropospheric contamination into the sample air, can be approximated with:

$$\delta_m \approx \delta_s f_s + \delta_c f_c \quad (14)$$

where δ_m is the final isotopic composition, and δ_s and δ_c represent the isotope composition of source and contamination and f_s and f_c are the fractional contributions to the total number of molecules after mixing. We further define the bias of the isotopic composition measurement as

$$\Delta\delta = \delta_s - \delta_m \quad (15)$$

Combining Eq. (8) and Eq. (9), we derive

$$\Delta\delta = (\delta_s - \delta_c) f_c, \quad (16)$$

For the calculation, we regard the mean differences between AirCore and LISA measurements (Table 5, e.g. 0.84 ppm for CO_2 and 1.8 ppb for CH_4) as the upper limit of bias induced in the stratospheric samples. Another estimate is performed based on the storage test results, that showed maximum drift of 0.11 ppm CO_2 and 2 ppb CH_4 , presented in Table 3.

The fraction f_c can be calculated according to Eq. (12) with C_s being the typical stratospheric mole fraction, which is taken to 395 ppm for CO₂ and 500 ppb for CH₄. We use typical tropospheric values of 405 ppm for CO₂ and 1800 ppb for methane. The isotopic compositions δ_c and δ_s are taken from various references, and are presented in Table 7. The resulting bias in the isotopic composition measurements are presented in Table 7.

We can readily see that the estimated uncertainties due to the storage bias are relatively small compared to the typical analytical precisions, also presented in Table 7. Hence the LISA sampler provides a viable sampling tool for accurate measurements of stable isotopes in CO₂ and CH₄.

Table 7: Expected bias in stable isotope measurements on samples obtained by LISA, due to the limited accuracy of the LISA sampler. Typical values for the troposphere and stratosphere are taken from the indicated references: A) (Troler et al., 1996) B) (Mrozek et al., 2016) C) (Nisbet et al., 2016) D) (Bergamaschi et al., 2001) E) (Aoki et al., 2003) and F) (Röckmann et al., 2011). Reported measurement reproducibility's, Re , for stratospheric air are also provided. $\delta^{13}C$ and $\delta^{18}O$ values are with respect to Vienna Pee Dee Belemnite (VPDB) and $\Delta^{17}O$ and δ^2H are with respect to Vienna Standard Mean Ocean Water (VSMOW). f_c was calculated using a source value 395 ppm (CO₂) and 500 ppb (CH₄). For f_{c1} contamination values of 0.84 ppm (CO₂) and 1.8 ppb (CH₄) based on LISA AirCore observed mean bias, resulting in $\Delta\delta_1$; For f_{c2} the maximum observed drift (Figure 2) of 0.11 ppm (CO₂) and 2 ppb (CH₄) are used, resulting in $\Delta\delta_2$.

Species	δ_c (‰)	δ_s (‰)	Re (‰)	LISA-AirCore		Storage Test	
				f_{c1}	$ \Delta\delta_1 $ (‰)	f_{c2}	$ \Delta\delta_2 $ (‰)
$\delta^{13}C$ (CO ₂) (VPDB)	-7.5 (A)	-8.4 (E)	0.02 (E)		0.077		0.01
$\delta^{18}O$ (CO ₂) (VPDB)	-2 (A)	12 (E)	0.05 (E)	0.086	1.203	0.011	0.158
$\Delta^{17}O$ (CO ₂) (VSMOW)	0 (B)	7 (B)	0.2 (B)		0.602		0.079
$\delta^{13}C$ (CH ₄) (VPDB)	-47 (C)	-20 (F)	0.7 (F)	0.006	0.134	0.006	0.149
δ^2H (CH ₄) (VSMOW)	-85 (D)	190 (F)	2.3 (F)		1.366		1.517

This might actually also explain the differences observed in April flights with respect to the September flights. During Fall NH CO₂ is expected to be much smaller, thus the difference between contaminant and actual stratospheric mixing ratio is much lower than during spring (when tropospheric CO₂ may be up to 10 ppm higher). This is also in line with the much higher deviations found in the laboratory experiments when having larger concentration differences.

A 0.8 ppm difference was observed during the storage test (Fig. 2) when a rough 250-ppm (assuming room air of about 405 ppm) concentration difference existed across the bags material over 4 hours. According to this storage test result, a seasonal cycle of 10 ppm in CO₂ between winter and summer would cause a difference of $10/250 \times 0.8 = 0.032$. Hence the observed increase of 1 ppm observed from the AirCore LISA difference cannot be fully attributed to the seasonal variability. This is now explicitly mentioned in the revised manuscript.

“The seasonality in tropospheric CO₂ that causes the difference between sampled air and its storage environment could contribute to the observed difference. The northern hemisphere winter CO₂ mole fractions are typically 10 ppm higher those in summer. During the storage test with low mole fractions, e.g. sample nos. 6&7 in Figure 2, a drift of up to 0.8 ppm was observed. Therefore, a typical seasonal difference of 10 ppm could only explain a difference of 0.03 ppm in the observed CO₂ bias.”

We would like to note that citing a publication in preparation, e.g. Chen et al. in prep, is in line with the guidelines of AMT. Nevertheless, we have added a sentence to provide more details about the AirCore measurements.

“The AirCore used during the campaign consists of two pieces of stainless steel tubing (40 m long ¼ in. OD and 60 m long 1/8 in. OD, wall thickness 0.01 in.), with a total weight of ~3.6 kg.”

Minor/specific observations:

p.3. l. 1: the Engel et al. trend is only representative for the mid-latitudes of the Northern Hemisphere, above 24 km altitude.

We have added this in the revised version.

“...however no significant change in the strength of the BDC in the northern hemisphere at mid latitudes was detected (Engel et al., 2009, 2017).”

p.3. l. 3.: A reference from 1983 may not be very good to point to current deficits in GCMs.

We have added a more recent reference (Gerber et al., 2012).

“In spite of all the efforts to make observations of stratospheric tracers, GCM’s remain poorly constrained (Gerber et al., 2012), a problem already pointed out several decades ago (Ehhalt et al., 1983).”

p.3. l. 26.: could you be more specific on the allowed weight?

The total payload of a weather balloon typically ranges between 0.2-12 kg, which represents the range of the weight from radiosondes to medium-weight scientific instrumentation. We have changed the original sentence in the revised version as

“... typically ranges between 0.2-12 kg...”

p. 3. L. 29.: as LISA samples during ascent, has possible outgassing of CO₂ from the Styrofoam be considered?

Yes, this has been considered. Currently, we don’t have data on outgassing under stratospheric conditions.

“Another potential source for the bias in the CO₂ mole fractions is outgassing from the packaging material and balloon. As the balloon ascends the surrounding pressure decreases and gasses desorb from the surface of the packaging material and balloon, which potentially influences the mole fractions. This would, however, not explain the good agreement during the September flights opposed to the April flights. Furthermore, the inlet is located at the top of the payload and any outgassing is flushed away from the inlet during ascent.”

p. 5. section 3.1. Could you comment on how dry the test samples were? This will make a very large difference, especially for CO₂.

Test samples were prepared with air from cylinders that typically contain less than 0.03% H₂O. This value has been added to the revised version. The direct measurements yielded values

below 0.06 % for the MLF bags and 0.12 % for the Tedlar bags, but since we used water vapour correction functions we did not add the absolute water content to the manuscript.

“...with dry air (<0.03% H₂O) from a cylinder”

p.6. l. 13.: How were the bags preconditioned. Why was N₂ used and not ambient air?

A small discussion on the preconditioning procedure has been added in the revised version. In principle, ambient air could be used as well.

“The bags were filled with N₂ from a cylinder and subsequently evacuated with a vacuum pump, prior to filling with test sample. In principle ambient air can be used to flush the bags, as long as it is dry.”

p. 7. Section 4.: At which temperature was the flow characterized? Strat. Temperatures are much lower, which may influence pump performance quite strongly.

Indeed, no laboratory tests were performed at stratospheric temperatures. The pump performance could be affected by the cold environment. First, the batteries could lose capacity and cause the power supplied to the pump decreasing. The temperature inside the thermally insulated package, where the pump is located, during flight ranged between 30 and -15 °C. Secondly, the diaphragm is exposed to the cold air passing through the pump. The elasticity of a rubber is temperature dependent, which could reduce the performance of the pump. On the other hand, heat is released during operation of the pump, which increases the temperature. Finally, the effect of air temperature on sample size follows the ideal gas law, and the sample size increases at low temperatures. As no experimental data is available to determine the performance of the pump at stratospheric temperature, we assume that the pump performs the same during flight as at room temperature for the calculation of the sample size.

We have clarified in the manuscript with the following text in section 6

“During the experiments described in Section 4, the pump was at room temperature. The pump performance could be affected by the cold environment. First, the batteries could lose capacity and cause the power supplied to the pump to decrease. The temperature inside the thermally insulated package, where the pump is located, during flight ranged between 30 and -15 °C. Secondly, the diaphragm is exposed to the cold air passing through the pump. The elasticity of a rubber is temperature dependent, which could reduce the performance of the pump. On the other hand, heat is released during operation of the pump, which increases the temperature. Finally, the effect of air temperature on sample size follows the ideal gas law, and the sample size increases at low temperatures. As no experimental data is available to determine the performance of the pump at stratospheric temperature, we assume that the pump performs the same during flight as at room temperature for the calculation of the sample size.”

p. 7. L 8 sample, not samples

The typo has been corrected in the revised version.

“...of air sample collected into...”

p. 7. L. 11 (and elsewhere in the manuscript): please be consistent in the use of L vs. L STP.

We have made sure that the manuscript is now consistent with the use of L at STP, except for the volume of an object, e.g. the volume of the bag is 2.58 litre.

p.8. l. 25: please restrict this to 30 hPa, as the flow was not measured at lower pressures.

We now clarify in the figure caption that the presented curves are modelled based on the fitted parameters. The extrapolated curves are informative, although they may contain larger uncertainties than those within the tested range. Therefore, we would like to keep the curves outside the tested range, but have made the point clear in the revised version.

“...Atmospheric pressure on the left and corresponding altitude on the right, as a function of modelled sample volume. Note that for atmospheric pressure >120 and <30 hPa, as well as for sampling time >150 s the modelled results rely on extrapolation of the observations.”

p. 9. L. 6.: Is burst pressure of the bag temperature dependent?

No experiments have been performed to show whether the burst pressure of the bag is temperature dependent. All bags survived the four flights. Two bags burst at room temperature at the pressure gradients across the bags of 312 and 331 hPa, respectively. The burst temperature might be temperature dependent; however, we think that the variation of the quality of the bags explains the different burst pressures at room temperature.

p. 9. L. 10: 240 K is actually very warm for the stratosphere.

Changed to 220 K, which is in line with the standard atmosphere.

“...cold e.g. 220 K, the total sampled volume at”

p. 12. L. 17 l. 27: please be consistent in using only one value for the deviations. If the value of 0.84 ppm is used, this is much larger than stated in Engel et al., 2017. Have the authors made an uncertainty estimate for the flight on April 26, considering that no AirCore coils temperature is available?

In the manuscript we have made sure that it is clear which deviation are discussed. The effect of coil temperature on the profile is relatively small.

p. 12 l. 26: It is not only molecular diffusion.

Indeed, we have revised the sentence to include also the effects from Taylor dispersion and air mixing in the cavity of the CRDS analyser. We also state that the altitude registration in pressure coordinates is sensitive to user input.

“Two aspects contribute to errors in the AirCore weighted mean. First of all, the AirCore profile needs to be weighted, since not all the pressure levels contribute equally to the sampler samples. The uncertainty in altitude of the AirCore profile adds a level of uncertainty to the AirCore weighted mean. Secondly, the retrieved AirCore profiles are already smoothed due to molecular diffusion and Taylor dispersion, and smearing effects in sample renewal of the cavity of the CRDS. Moreover, the AirCore profile suffers from uncertainty in altitude registration.

For more information on the uncertainties associated with AirCore profiles we refer to Engel et al.(2017), Karion et al. (2010) and Membrive et al. (2017)."

p. 12/13. Section 6.1.: This whole section lacks a conclusion. First, which is the best estimate of the deviations, and second what could cause the strong deviations during the April flight (see also my major comment above). Have pollution/outgassing been considered to explain the differences?

A more elaborate discussion has been provided. Outgassing was considered, and we have added this to the discussion.

"Another potential source for the bias in the 26 April 2017 flight is outgassing from the packaging material and balloon. As the balloon ascends the surrounding pressure decreases and gasses desorb from the surface of the packaging material and balloon, which potentially influences our measurements. This would, however, not explain the good agreement during the September flights opposed to the April flights. Furthermore, the inlet is located at the top of the payload and any outgassing is flushed away from the inlet during ascent."

The suggestion of the influence NH seasonality, as in the major comment, has been accounted for.

"The seasonality in tropospheric CO₂ could be another explanation to the observed difference. In northern hemisphere winter CO₂ mole fractions are higher, and hence the difference between sampled air and storage environment is higher. However, during the storage tests with low mole fractions, e.g. sample nos. 6&7 in Figure 2, the mole fractions of CO₂ were much lower than typical stratospheric samples. In that experiment the drift did not exceed 0.8 ppm. Hence, such a large drift cannot be fully attributed to the CO₂ variability in the storage environment that is due to seasonality."

A more detailed discussion on the uncertainties in AirCore has been added, see also the reply to the previous comment. We have no good reason to say that either one of the estimates is better, other than the faulty datalogger, therefore we conclude with deviations as follows:

"Even though the AirCore might have been affected by diffusion and problems with the datalogger, AirCore and Sampler show good agreement in all flights, with mean AirCore sampler differences of 0.84 ppm for CO₂, 1.8 ppb for CH₄ and 6.3 ppb for CO."

p. 13. L. 14.: this sentence is odd. How should a prolonged sampling time result in increased vertical resolution?

The sentence has been rephrased. The explanation why prolonged sampling time compromises vertical resolution was presented in section 4.

"Therefore, the gain in sample amount, for example adding 10 extra seconds of sampling time, is small; however, the decrease in vertical resolution is significant"

p. 14.l. 7.: this sentence is wrong, uncertainty is mentioned twice.

The sentence has been corrected.

“The uncertainty in sounding of the pressure sensors is 1 hPa at 200 hPa (Vaisala, 2013). The total uncertainty after 200 seconds of sampling is 9 mL_{stp}, slightly higher than the effect found above.”

p. 14. Section 6.4. see major comment above. I believe that the contaminating fraction may be much higher.

See our reply to the major comment.

p. 15. L. 24: this is not up to 0.84 ppm. 0.84 ppm is actually an average deviation if all samples are considered.

Correct, we have rephrased the sentence and now use “mean difference” instead of “up to”.

“...showing the mean differences between AirCore and LISA of ...”

Figure 2: the colours for the different bags types are virtually undistinguishable.

The colours of all figures have changed. With respect to Figure 2, we have also increased the marker size to improve visibility.

Figure 3: I wonder if all panels are needed here. I suggest removing panel b.

Panel b shows the linear behaviour observed between sampled volume and ambient pressure. This allows us to model the volume sampled according to Eq. (2) where $a(t)$ is then the only parameter depending on sampling time. To justify our model, we deem it necessary to show panel b. Panel c shows the fit that models the time dependence of the linear coefficient $a(t)$ according to Eq. (3).

Table 6: please explain the different columns on the table heading, not only in the text.

We have provided more detailed explanations of the different columns in the revised version, now it is updated to Table 7:

“Table 7: Expected bias in stable isotope measurements on samples obtained by LISA, due to the limited accuracy of the LISA sampler. Typical values for the troposphere and stratosphere are taken from the indicated references: A) (Troler et al., 1996) B) (Mrozek et al., 2016) C) (Nisbet et al., 2016) D) (Bergamaschi et al., 2001) E) (Aoki et al., 2003) and F) (Röckmann et al., 2011). Reported measurement reproducibility's, R_e , for stratospheric air are also provided. $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ values are with respect to Vienna Pee Dee Belemnite (VPDB) and $\Delta^{17}\text{O}$ and $\delta^2\text{H}$ are with respect to Vienna Standard Mean Ocean Water (VSMOW). f_c was calculated using a source value 395 ppm (CO_2) and 500 ppb (CH_4). For f_{c1} contamination values of 0.84 ppm (CO_2) and 1.8 ppb (CH_4) based on LISA AirCore observed mean bias, resulting in $\Delta\delta_1$; For f_{c2} the maximum observed drift (Figure 2) of 0.11 ppm (CO_2) and 2 ppb (CH_4) are used, resulting in $\Delta\delta_2$.”

References

Gerber, E. P., Butler, A., Calvo, N., Charlton-Perez, A., Giorgetta, M., Manzini, E., Perlwitz, J., Polvani, L. M., Sassi, F., Scaife, A. A., Shaw, T. A., Son, S. W. and Watanabe, S.: Assessing and understanding the impact of stratospheric dynamics and variability on the earth system, *Bull. Am. Meteorol. Soc.*, 93(6), 845–859, doi:10.1175/BAMS-D-11-00145.1, 2012.