The authors would like to thank the reviewers for their valuable comments which helped in improving the quality of the manuscript. Our point-by-point responses to the reviewer's comments appear in bold below.

Responses to Review #1

If the authors wanted to improve the paper they could analyse the bias between lidar to IASC / ERA5 indepentently for the dry and wet period and seperated for below 1.5km altitude and above 3 km altitude.

This comparison between the two periods is already discussed in detail in the manuscript. It is presented in Figure 5 and Table 4 of Section 4.2. The results for the <1.5 km and [1.5–3] km altitude ranges are provided in Table 4 (columns 3 and 4). To make this point clearer, we have added the phrase "for both the wet and dry periods" to the caption of Table 4.

Responses to Review #2

Major Comments:

1. Concerning my previous major comment 3 that related to the method of normalizing the lidar calibrations through comparison with ERA5. It is very appealing to use satellite retrievals for calibration normalization in a study like this where an intercomparison of instruments was not done prior to the measurement campaign. To my knowledge this application of satellite measurements as a kind of transfer standard has not been done previously. For this reason, this manuscript could well become a seminal reference for performing such analysis in the future. It is therefore very important to clearly justify the technique used. As it stands, I still do not see how it can be justified to use only part of the regression analysis (the slope and not the intercept) between lidar and ERA5 for the calibration adjustment.

In their response authors state: "As we explained in our response to Reviewer 1, the lidar- ERA5 mean bias is highly correlated with the intercept. We decided to correct only the slope using ERA5 to avoid differences in the way all lidars were calibrated."

I really do not follow this logic. Correcting only the slope does not "avoid differences in the way all lidars were calibrated". Instead, by not utilizing the full regression equation to correct each lidar it makes the technique for calibration in fact different for each system since a differing amount of bias is unaccounted for in each separate lidar calibration adjustment.

Authors state "the offset remained uncorrected in order to maintain natural differences in the data independent of calibration." How does not making an offset adjustment maintain natural differences in the data while making the slope adjustment would presumably not do so? Both slope and offset come from the same regression. I don't see how they can be validly separated in a calibration adjustment.

1. This is a critical point to rigorously justify and I am not persuaded by the current arguments. What does the calibration adjustment look like if the full regression equations are used or if the mean/median ratio of ordered pairs used as described in my previous comments?

2. To fully justify this technique so that it is able to be referenced in the future the authors really must spend more time critically studying this calibration adjustment. Please show the reader the effect of performing the full calibration adjustment using slope and intercept (or mean/median ratio of calibration constants) and compare with what the authors prefer to do

(only use the slope) and argue more persuasively why this partial calibration adjustment is acceptable to do.

We understand the reviewer's request to assess the impact of correcting the intercept on the reported results. However, the intercept represents a systematic bias in the data. Correcting for it would have no effect on the COV or RMSE metrics; only the mean bias would be shifted by the value of the intercept. This adjustment is straightforward to calculate, and we believe that presenting these results separately would add little value. For this reason, we do not consider it appropriate to duplicate our tables and figures to include an additional set of results based on intercept-corrected data.

Nevertheless, we agree with the reviewer that the explanation should be more specific. We have therefore added the following paragraph:

"ERA5 data may nevertheless be biased. Biases in ERA5 arise from various sources: (i) inadequate accounting for instrumental errors during the data assimilation process; (ii) assumptions involved in the simulation of satellite radiances; (iii) deficiencies in the parameterizations of convection, evaporation, and turbulent exchanges; (iv) limited spatial resolution; (v) imperfections in the representation of the planetary boundary layer (PBL); and (vi) uncertainties in the representation of clouds and precipitation. Unlike ERA5 data, after correction for the baseline and sky radiance, the ratio between the water vapor and dinitrogen Raman channels tends toward zero in the absence of atmospheric water vapor. For this reason, the lidar calibration relies solely on the slope derived from the comparison between the WVMRs retrieved from the Raman lidar and those from ERA5. The intercept is therefore assumed to arise exclusively from biases in the reanalysis data."

We have also added: "Note that, in the absence of atmospheric water vapor, no Raman backscattering is produced by water molecules, and the signal recorded on the H_2O Raman channel corresponds solely to the background noise associated with the acquisition baseline and sky radiance. These two contributions are subtracted from the signal for each lidar used in this study."

If the intercept is included in the calibration, it will introduce a systematic bias that originates exclusively from the model. This bias will be added to the biases calculated as a result of this study. It should be emphasized that we are not dealing directly with satellite observations for the calibration, but rather with model outputs. Satellite measurements are assimilated into the reanalysis system, which relies on a radiative transfer model to simulate satellite radiances.

Minor Comments

The paper is considerably more detailed on the issues of biases due to particle differential transmission and the temperature sensitivity of Raman scattering. Thank you for this considerable amount of effort; I believe the community will benefit.

1. Please include a clear statement that both of these biases lead to an apparent moistening of the water vapor profile and that these biases add together.

This explanation has been added in subsection 3.1: "All these biases contribute to an apparent moistening of the water vapor profile and are cumulative."

1. It should also be stated that this is an explanation for some of the difference between lidar and satellite noted in the manuscript so please also quantify the magnitude of these combined biases using similar units. Currently the bias between lidar and satellite is generally given in terms of g kg-1 while these biases are expressed in percentage. It would aid the reader to have, somewhere in the manuscript, the magnitude of these uncorrected biases expressed in g kg-1.

Since the reviewer rightly requested that the mean WVMR profiles also be provided, it is straightforward to convert the relative biases expressed in percentages into grams per kilogram. In Appendix B, we use percentages because the biases affect specific factors and therefore influence both low and high water vapor contents in the same manner.

2. Table B1 is very helpful. One can do the math and see that for RL2 and RL8 the combined effect of these neglected biases might reach almost 7% depending on their separate height dependence. Please also convert these biases to g kg-1 in a representative example so that comparison with the rest of the paper is easier to do.

In Table B1, we have added rows showing the altitude ranges for which the biases are given in g kg^{-1} for representative values of WVMR.

Table B1: Biases on the water vapor mixing ratio (WVMR) for each Raman lidar due to i) the temperature sensitivity of the backscatter cross section of the H_2O-Raman channel and ii) the differential transmission of aerosols. Two aerosol loads are considered. The first has an aerosol optical thickness (AOT) of 0.1 and an Ångström exponent of 1.5. The second corresponds to the presence of Saharan dust layers, with an AOT of 0.5 and an Ångström exponent of 0.5.

		RL1	RL2	RL3	RL4	RL5	RL6	RL7	RL8
Emitted wavelength (nm)		~354.7	354.69	354.71	354.83	354.72	354.74	354.74	354.70
H_2O-Raman channel									
Central wavelength of filter (CWL) (nm)		407.46	407.56	407.65	407.62	407.50	407.54	407.58	407.5
Filter bandwidth (FWHM) (nm)		1.06	0.30	0.35	0.37	0.25	0.30	0.35	0.30
Bias									
Bias due to aerosols AOT = 0.1 (PBL)		< 0.7%	< 2%	< 0.7%					< 2%
Bias due to aerosols AOT = 0.5 (with dusts)	≤ 1.5 km	< 0.5%	< 1%	< 0.5 %					< 1%
	> 1.5 km	< 1.3%	< 4.7%	< 1.3%					< 4.7 %
Bias due to temperature (ΔT = 30 K)		< 0.3%	2.70%	2.7%	0.5%	2%	1.8%	2.8%	2.2%
Total maximum bias	≤ 1.5 km	< 1.2%	< 4.7%	< 3.4%	< 1.2%	< 2.7%	< 2.5%	< 3.5%	< 4.2%
	> 1.5 km	< 1.6%	< 7.4%	< 4%	< 1.8%	< 3.3%	< 3.1%	< 4.1%	< 6.9%
Bias for representative values of WVMR (g kg^{-1})									
WVMR = 7.5 g kg^{-1}]0.2 1.5] km		< 0.09	< 0.35	< 0.26	< 0.09	< 0.20	< 0.19	< 0.26	< 0.32
WVMR = 5 g kg^{-1}]1.5 3] km		< 0.08	< 0.37	< 0.20	< 0.09	< 0.17	< 0.16	< 0.21	< 0.35
WVMR = 2.5 g kg^{-1}]3 5] km		< 0.04	< 0.19	< 0.10	< 0.05	< 0.08	< 0.08	< 0.10	< 0.17

3. Figure B1 is quite helpful. Please clarify which RL this plot pertains to. Or does this figure somehow present an average effect for all RLs? If this figure is somehow an average, the figure would be aided by plotting bands that show the range of biases for all lidar systems. Somehow the full range of the bias should be represented.

Yes, this is an average effect. These biases do not depend on the instruments; rather, they are related to the amount and type of aerosols present in the atmosphere. The inversion algorithms used for the different lidars are identical, and the resulting correction curves are therefore the same. The two aerosol-loading scenarios considered are representative of all sites:

"During the WaLiNeAs field campaign, we observed two contrasting scenarios regarding aerosol loading in the lower and middle troposphere, as recorded by all the lidars."

1. Please add to this figure the already calculated temperature dependence as a function of altitude. This need only be estimated but it would be very helpful to show the combined effects of the two biases.

Table B1 now includes the sum of the bias contributions, reported as the total maximum bias.

In addition, we have included an example of the temperature-dependent bias profiles for each lidar site and added the corresponding figure: "As an example, the vertical profiles of the maximum bias for each lidar site are shown in Figure B2. Such a bias magnitude is mainly related to the filter bandwidth."

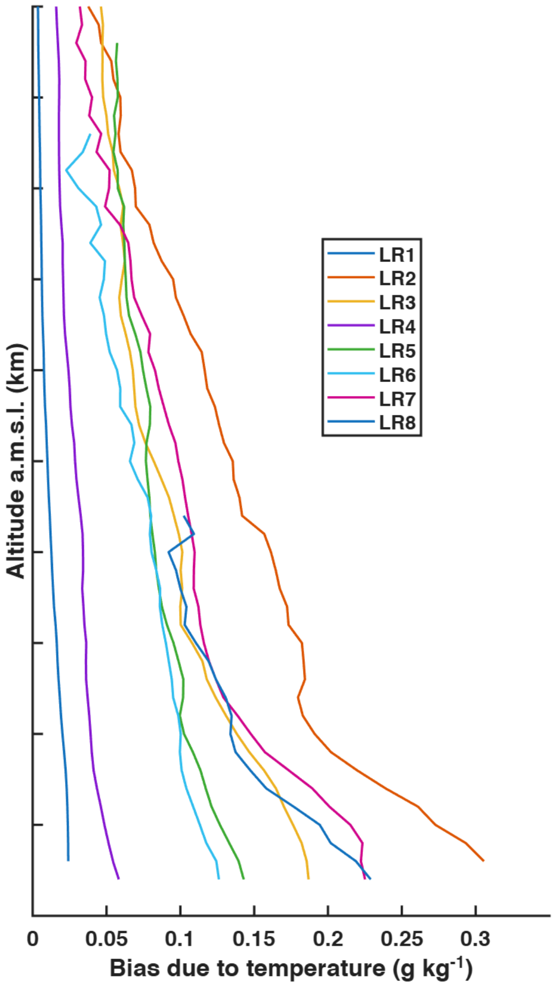

Figure B2: Vertical profiles of the maximum bias for each lidar site (LR1 to LR8, see Table 1). The considered vertical profile of the water vapor mixing ratio is the one on Figure 4c.