Reply to editor, AMT-2015-63, Barrera-Verdejo et al.

First of all, we would like to thank Professor Sica for his detailed revision of our manuscript. In the following you will find our replies to your concerns: the grey color represents the review, in black our comments and in italic the text of the manuscript. The line numbers refer to the first submitted manuscript version.

Review of AMT–2015–63, Barrera-Verdejo et al

1. This paper introduces an interesting new technique, which could improve our ability to combine and enhance our measurements of water vapour and temperature by radiometer and lidar. However, before posting this manuscript to the AMTD site I have a few comments I recommend you address, as when referees are assigned in the second phase of the review I anticipate they will have similar comments.

The first comment is a technical one.

Anchored Note, page 8

> *Therefore, before the analysis, a running average is performed on the data (we choose 300 m window size) previously tobefore the analysis.*

The data is related to the forward model via

$$\mathbf{y} = \mathbf{F}(\mathbf{x}, \mathbf{b}) + \varepsilon$$

The measurement is on the measurement grid while the retrieval parameters must be interpolated from the retrieval grid (as an aside, did you mention what kind of interpolation you use, if not you need to).

There is no problem co-adding your measurements at 30 m to make a new data grid at 300 m, from which you can retrieve on any resolution retrieval grid you wish. However, using a smoothing on the measurements in the retrieval may introduce off-diagonal elements to the measurements covariance matrix ($\mathbf{S}_\mathbf{y}$). This complication may explain why many of your nights don't process for water vapour, you may have affected the noise (maybe by de-whitening it).

We probably should have made clearer that the running average is applied only to estimate the maximum altitude where the lidar data is considered valid and that it is only applied to the error profile. As explained in the manuscript, this running average is needed to avoid possible peaks in the lidar error, which would lead to an erroneous identification of a trustworthy lidar range. It is written in the manuscript (from line 244) as follows:

> *For every lidar profile one must determine the range of altitudes where the data can be considered meaningful. This range has been defined via the relative error. The relative error is calculated at each altitude as the ratio between the error and the measurement, as a percentage. When this value is larger than 100%, the data is considered too noisy and is discarded. Care is needed when defining this threshold, because possible random peaks in the error can lead to a misidentification. Therefore, before the analysis, a running average is performed on the data (we choose 300 m window size) previously before to the analysis. In general, the 100% error altitude might be reached at different points depending on the weather situation or night/day-times.*

The information required in the optimal estimation equation (observation vector **y** and matrix $\mathbf{S_e}$) is introduced straightforwardly, without any smoothing.

There are indeed some night periods that were not processed during HOPE. But this has nothing to do with convergence issues; they were not processed due to the presence of frequent nighttime clouds, where our cloud-free algorithm is not applicable.

2. The second reservation concerns the lack of information available to the reader to judge the quality of the retrievals. For example:

Anchored Note, page 11

> In the following, the results for one example profile are presented (Fig. 2).

For this figure (and subsequent ones) to be meaningful we need to see the following:

- the Jacobians
- the averaging kernels
- the vertical resolution with height
- the residuals (as well as the cost)

The above holds true for each one of your retrievals, and also when you show a table with degrees of freedom: we need to see the averaging kernels, particularly as some of the numbers in the tables appear to be different from what one may have guessed.

We agree that it is useful to provide additional information. To help the reader interpret the tables showing the degrees of freedom for signal, we have now included in the manuscript an additional plot (new Fig. 8) showing the vertical profiles of cumulative degrees of freedom for signal, derived from the diagonal of the averaging kernel matrix. This plot has been commented accordingly in section 4.2.5, where the DOF where initially introduced.

Apart from for the sake of brevity, we argue against showing plots of Jacobian matrices, residuals and/or complete averaging kernel matrices in the manuscript; from an expert's point of view they could be interesting to analyze, but for a general reader additional graphs would distract from the main statements of our manuscript. We believe that the proposed assessment based on the theoretical error and degrees of freedom for signal provides a thorough insight into retrieval performance. delivers us a complete enough picture. Second, for the sake of brevity, because the manuscript is already quite long and contains a considerable number of plots.

For your information, we have included some example plots and an explanation of a water vapour profile (see Figures in the Appendix). We would be happy to include these in a supplementary dataset but we feel that showing and interpreting these in the main body of text would lead to an unnecessary complication of the matter.

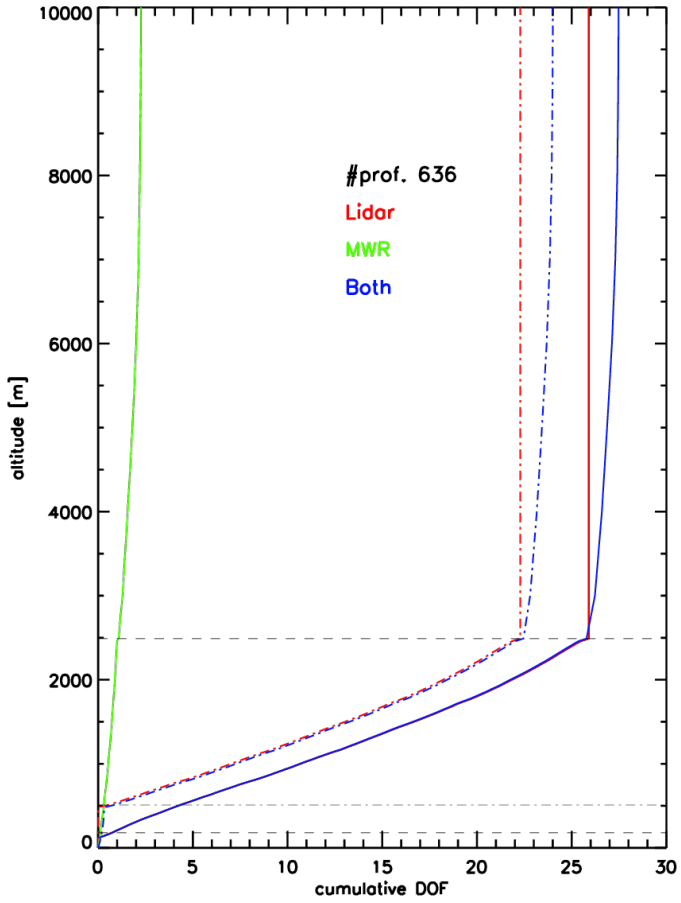

Figure 1: (Fig. 8 in the manuscript): Cumulative degrees of freedom per profile for the different instrument combinations: in red, only-RL; in green, only-MWR and in blue, the combination of the two sensors. The dotted-dashed lines represent the degrees of freedom for the case where the overlapping function has been extended up to 500 m. The average numbers of DOF in each region are summarized in Table 2. The dashed horizontal grey lines enclose the part of the atmosphere where lidar data has been considered.

3. I am also concerned about the use of "theoretical" (as used in the paper) rather than actual uncertainties, that is the uncertainties one can estimate to test the theoretical uncertainty model. It is misleading to suggest your method improves the uncertainties of real measurement combinations when you do not use the actual uncertainties of the measurements as a comparison, particularly as their are both systematic and random uncertainties.

During the development phase of any retrieval, such kind of error estimation (reported in literature as theoretical, theoretically estimated, a posteriori error, etc.) is of capital importance for investigating retrieval results. It is an essential parameter for evaluating the meaningfulness of the two instruments synergy.

We make clear that our method improves the uncertainties of real measurement combinations, of course assuming that the initial uncertainty estimates ($\mathbf{S_a}$, $\mathbf{S_e}$) are properly defined. In this case the theoretical error would correspond to the actual retrieval error. Nevertheless, this initial error definition is quite sensitive and, in any

case, based on our best knowledge. This is now mentioned in the manuscript in line Sec. 4.2.5.

The major problem we face is that no “truth” for water vapor exists and therefore we cannot establish the “actual uncertainty”. We choose to consider the radiosondes as the closest description of the true atmospheric state although they suffer from several problems (see below). Unfortunately, only few radiosondes are available for comparison purposes. A statistical comparison, based on longer time series, would be needed and will be the focus of further studies.

4. Anchored Note, page 16

launch distance of 4 km to the site, drifting of the balloon, dry bias, etc.

Actually these aren’t that hard to assess? For instance, isn’t there a barometer and maybe a flux tower near the lidar? You can compare that to the place the sonde was launched and see how similar or different it is. As far as the drift of the balloons, the balloons are tracked so you can easily tell if they are flying over you or not. Meteorological maps can help you assess if you are in the same air mass: maybe you want to start only with sounding where you think you are in the same air mass? As far as biases there are many references on radiosonde biases (e.g. Miloshevich and colleagues), please quote the relevant papers in the literature.

Atmospheric moisture is highly variable in the atmospheric boundary layer due to turbulence (see plot on the right). Therefore a radiosonde drifting along its trajectory encounters different humidity values from those measured by an upward looking instrument at a fixed position. As these water vapor changes act on second-to-minute scale, meteorological maps (from an NWP model?) do not help. Many studies address radiosonde biases for Vaisala but the radiosondes from Graw that we use contain an internal humidity correction by the manufacturer.

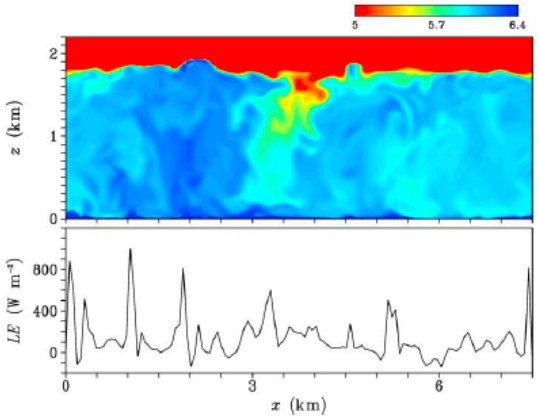

Figure 2: (Fig. 10 from Shao et al., 2013, Boundary layer Meteorology): Snap shot of a cross section of specific humidity and latent heat flux.

5. The sensitivity study suggests overlap is the primary uncertainty in the water vapour measurements. Overlap can’t be the problem since you are dividing count profiles. For an achromatic system the overlap is the same for each channel and cancels, hence why this technique is used despite being uncalibrated. However, there may be a differential overlap between the channels, but you should be able to estimate it and improve your correction (it has probably been studied in previous papers with the BASIL system, if not you can probably estimate by comparing the small and large telescopes (best) or from the comparison with the sondes (better than nothing)). Then the overlap could be a term in a realistic uncertainty budget. But perhaps you mean that due to the overlap the signal levels are low and thus, the statistical uncertainties are high. Please clarify this. I imagine the BASIL measurements have been well characterised and much of this information is

available.

Also, there is no discussion of aerosols, which in the PBL could (and likely do) have a much larger affect than the differential overlap. Perhaps during the campaign there are ancillary measurements of aerosols, or they can be estimated from the lidar measurements.

We interpret this point as the aerosol contribution to the differential transmission term (DT), which accounts for the different atmospheric transmission at the two Raman wavelengths of water vapour and molecular nitrogen. Said term is present in the algorithm used to obtain the water vapour mixing ratio from the Raman signals power ratio *q:*

q=(P_H2O/P_N2)*c DT,

with c being the calibration coefficient. The differential transmission has two components: one associated with Rayleigh (molecular) scattering and one associated with Mie (aerosol) scattering. The first one is by far the predominant, and can easily and precisely be computed based on the use of radiosonde or standard atmospheric profiles of number density. The second component, associated with the wavelength dependence of particle extinction, usually accounts for a very small portion of DT (1-2 %) and it is determinable from lidar measurements of particle extinction at 355 nm (Whiteman, 2003).

The molecular nitrogen Raman lidar signals at 387 nm (among other purposes) can also be used to determine the vertical profiles of the particle extinction coefficient at 355 nm. This profile is calculated based on the algorithm defined by Ansmann et al. (1992) and can be used to estimate the aerosol contribution to DT. We must emphasize that this correction accounts for a very small portion of DT (1-2 % also in case of high aerosol loading in the PBL). So, neglecting this term would imply a systematic error not exceeding 1-2 %, which is by far smaller than the random error affecting the water vapour mixing ratio measurements.

While the aerosol contribution to DT has been properly estimated for the purpose of these measurements, ignoring it (which is done by most scientists running water vapour Raman lidar systems) would have negligible effects on water vapour mixing ratio profiles from Raman lidars.

In addition, we also had in mind the reviewer's second motivation, i.e. that due to the overlap effect, the signal levels in the overlap region are smaller than those observed at higher levels and consequently the statistical uncertainties are larger than at higher levels.

6. Anchored Note, page 19

> *we artificially incremented the RL error by a factor of 4 to study the sensitivity of the retrieved profile with respect to the RL measurement uncertainty.*

There is no reason to expect the RL systematic uncertainties are 4 times that of the random ones. There is plenty of real information available to make reasoned systematic uncertainties estimates and which can then be compared to your theoretical estimates. You are trying to get results from real measurements, thus, you have to ultimately work with the actual uncertainties. That being said there is nothing wrong with "turning the knobs" on the covariances and exploring the robustness of the method, but the baseline should be using the experimentally determined uncertainties.

The aim of the exercise with inflated Lidar error is to show how the synergy would change in such a case, since we suspect that the given statistical lidar error is very small. The factor of four for the increment is not chosen randomly. We extract this factor from the comparison to the radiosonde profiles: we noticed that the mean deviation to the radiosondes at an altitude range of 1.5-2.5 km is four times larger than the theoretical error in this same region.

We do not expect any systematic error, since the optimal estimation method on which the presented retrieval is based, is built on the assumption of unbiased measurements with Gaussian random errors.

7. Please consider revising the paper to address these comments, as well as consider my other suggestions such as including more than 1 night for temperature so we know the one you picked wasn't the only one that worked. The introduction would benefit from some re-writing.

Thanks for the comments which led to the inclusion of an additional figure. In respect to the single case study for temperature: On the one hand, for the water vapor retrieval we include one time series afternoon as an illustrative example, even though we have retrieved more than 53 non-continuous hours of good water vapor profiles. On the other hand, for the temperature retrieval we present only one profile. This is because of the much reduced lidar temperature data availability, as explained in the manuscript (lines 482-484). Unfortunately, only four case studies have been processed for the lidar temperature, in which we can only find one very short clear sky interval close to a radiosonde ascent. And this is the case we present.

The retrieval of relative humidity, which is the important parameter for cloud formation, also requires the temperature profile. As the lidar temperature profile is only scarcely available we investigate the feasibility in retrieving the relative humidity profiles without using the lidar temperature.

It is clear that by presenting one single profile for temperature and relative humidity, the retrieval performance is not comprehensively analyzed. As you mention, it could happen that the presented example was the only profile working. Nevertheless, we are confident that this is not the case, because the temperature algorithm is based on the same scheme as the water vapor algorithm, which has been proved to work successfully for a long-time series. In summary, this paper is a feasibility study for the retrieval of relative humidity and longer time series will be assessed in future studies.

I am recommending "Publish subject to minor revisions (Editor review)." When you resubmit your revised manuscript please include a copy or detailed description which shows what changes you have made to address my concerns.

The changes on the manuscript are highlighted in red.

Best regards,

And on behalf of the co-authors,

María Barrera Verdejo
Susanne Crewell
Ulrich Löhnert
Emiliano Orlandi
Paolo Di Girolamo

Extra references

Ansmann, A., M. Riebesell, U. Wandinger, C. Weitkamp, E. Voss, W. Lahmann, and W. Michaelis, 1992: Combined Raman Elastic-Backscatter LIDAR for Vertical Profiling of Moisture, Aerosol Extinction, Backscatter, and LIDAR Ratio. Appl. Phys. B 55, 18-28, GKSS 92/E/52.

Shao., Y. S. Liu, S. Crewell, and J.H. Schween, 2013: Large-Eddy Atmosphere - Land Surface Modeling over Heterogeneous Surfaces: Model Development and Comparison with Measurements. Boundary Layer Meteorology, August 2013, 148 (2), 333-356, http://dx.doi.org/10.1007/s10546-013-9823-0.

Whiteman, D.N., 2003: Examination of the traditional Raman lidar technique. II. Evaluating the ratios for water vapor and aerosols. Appl. Opt. 42, 2593-2608.

Appendix:

1. Averaging kernel

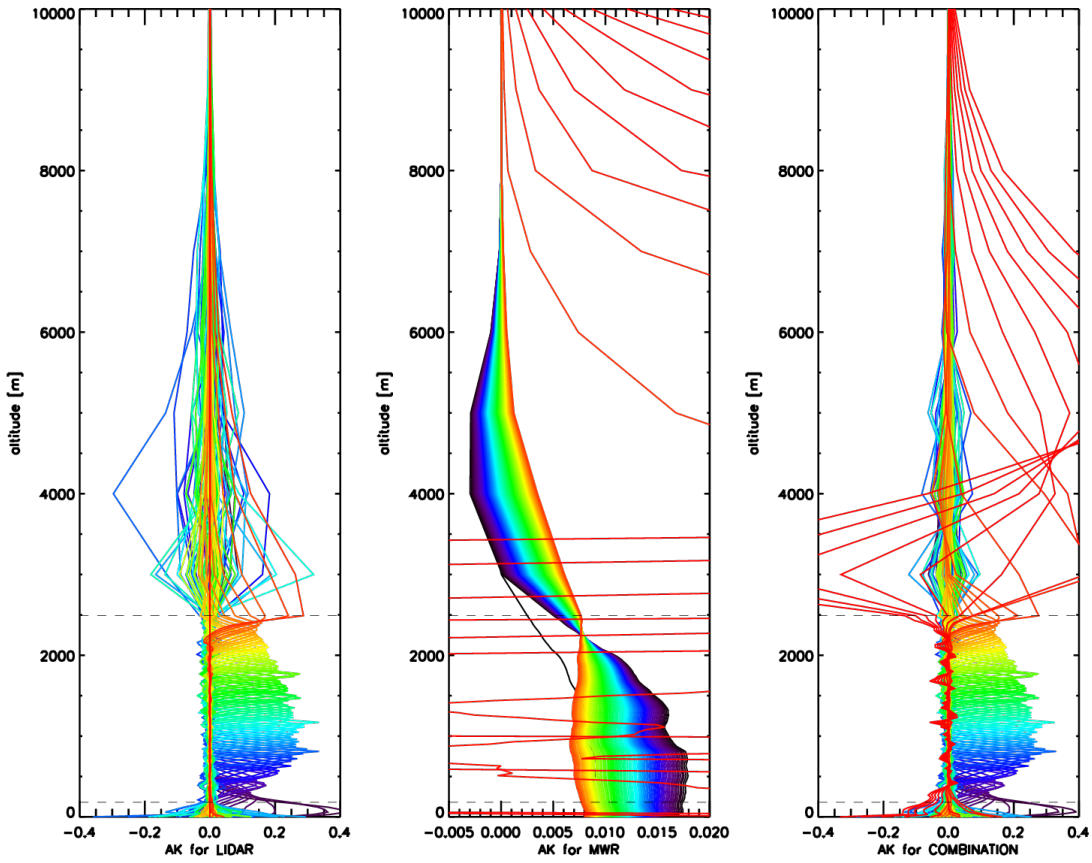

Figure 1: averaging kernels for (a) only-RL, (b) only-MWR and (c) the combination.

From left to right, the graphs are: the averaging kernel matrix from a) only lidar, b) only MWR and c) the combination of the both instruments. The color scale corresponds to the different altitudes: ground is represented by black, higher altitudes are represented with blue to red colors. For this profile, the lidar data is considered useful from 180 meters to 2.5 km, a region where the grid vertical resolution is 30 meters. When no lidar data is available, the vertical grid is reduced to 1 km.

The a priori correlation matrix plays an important role in the distribution of the water vapor information. It is important to note that the retrieval grid is not constant. Because of that, a perturbation on a thicker layer (i.e.: 1 km) produces a much higher variation on the retrieval. This variation is evident not only in higher layers, where the retrieval grid is coarser, but affects also lower layers. This is because the information in different atmospheric layer is not independent: the water vapor altitude correlation is defined by the a priori correlation matrix (Figure 1 in manuscript). For example, if the a priori correlation matrix was diagonal, we would see clearly weighing functions with the shape of a Dirac delta for every altitude of the lidar (see Figure 2 in Appendix).

For altitudes from ground to 2.5 km, the averaging kernel plot (Fig. 1a) shows narrow peaks at altitudes corresponding to Lidar measurement heights. Variations in higher layers are induced by the vertical correlations of the a priori q profile.

The only-MWR (Fig. 1b) provides much lower information content, as was already discussed in the manuscript. Please note that the scale in the figure is one order of magnitude smaller than for the lidar. The strong variation at 3 km shown by the red lines correspond to the variation of the retrieved profile due to a variation of the real profile at around 9 km. This strong lines are explained because of the non-uniform retrieval grid and the negative correlations between humidity at 9 and 3 km, as shown in figure1 of the manuscript.

For the combination of the two instruments the situation changes. For example, the strong waves (in red) induced by variations in high atmospheric layers, are strongly reduced in the lower atmosphere. This is because lidar data is considered in this region, and the lidar error is much smaller than the apriori uncertainty.

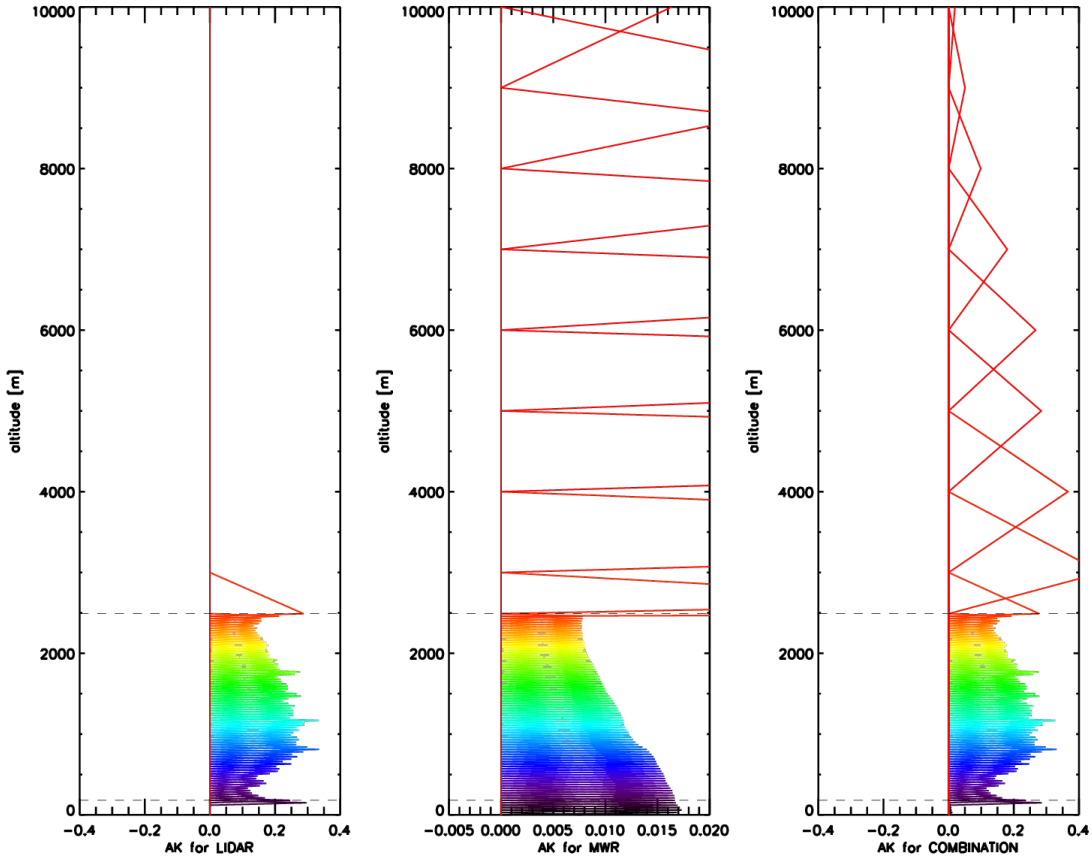

Figure 2: From left to right: averaging kernels of only-RL, only-MWR and combination, when using a diagonal a priori covariance matrix.

2. Jacobians

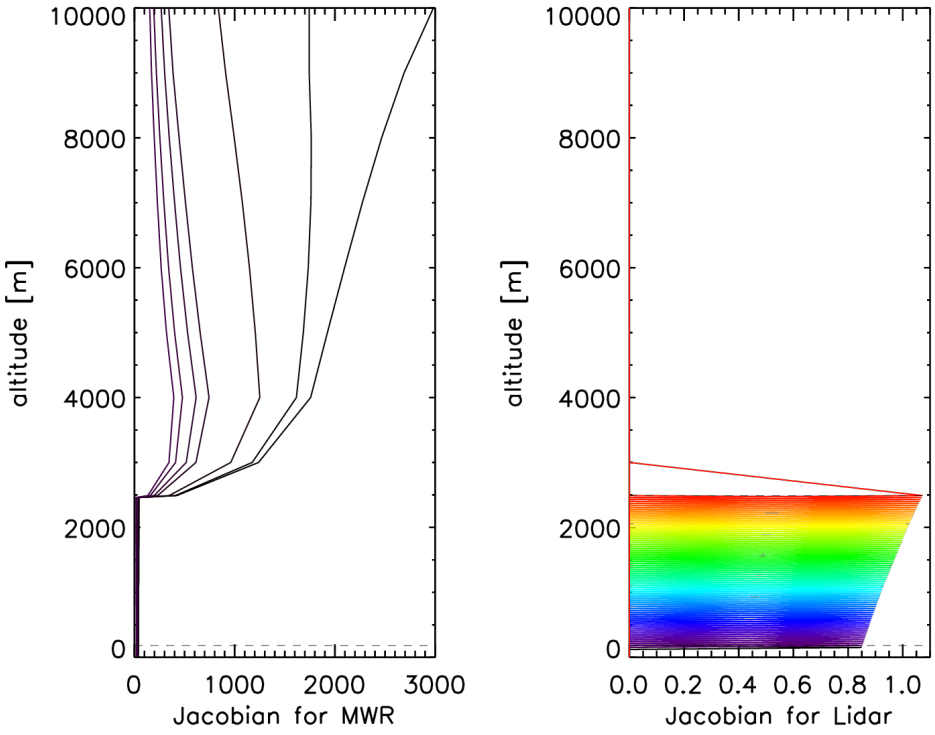

Figure 3: Jacobians for the MWR (left) and RL (right).

3. Vertical resolution

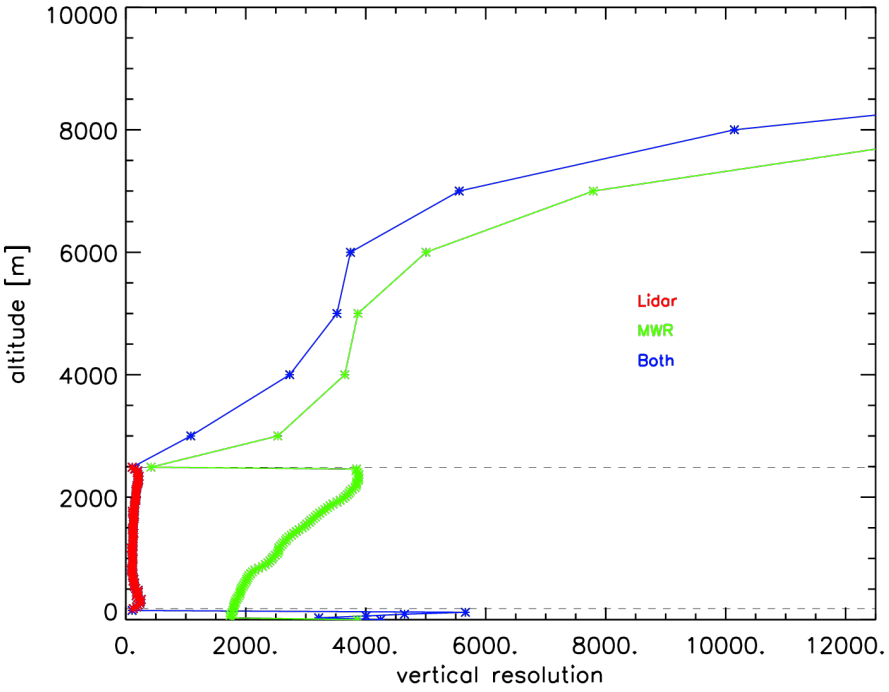

Figure 4: Vertical resolution

Replied to the comments on the manuscript

In addition to the file "Reply to editor.pdf", the present file and "ATM_barrera_red.pdf" are provided. In the present document, we reply the editor's comments on the manuscript. In order to ease further corrections, "ATM_barrera_red.pdf" provides a version of the manuscript where we highlight the changes on the text in red color.

The editor's comments have been copied from the manuscript and commented one by one. The number of pages and lines are referred to the original form of the manuscript. In grey, the editor's comments are presented, highlighted in italic are parts of the text the editor's comments refer to. In black, the answer from the authors can be read.

PAGE 1:

- Abstract The abstract reads poorly and needs significant editing.

Abstract edited in the manuscript.

- *In order to better understand these processes, highly resolved, accurate and continuous measurements of these parameters are required.* Measurements of humidity and temperature at high space? time? are required for the description of any meteorological event.

Both, high spatial and time resolutions are required. Even though typically time resolution is not an issue, finding a good compromise between continuous measurements and good vertical resolution is usually a big problem.

- *Unfortunately, instruments available nowadays are 5 not able to provide sufficient spatial resolution to describe short time scale processes.* I don't believe this unjustified statement is correct please remove it. Turbulence measurements can be at extremely high temporal-spatial resolution. Video images of clouds can have 1/30 of a sec temporal resolution.

Yes, that is true. But unfortunately, turbulence measurements are commonly restricted to instruments confined to (or at least close to) the surface and naturally, cloud cameras are not able to capture values of absolute humidity, relative humidity and temperature, which is our goal.

- *Optimal Estimation Scheme (OES).* please use more common name Optimal Estimation Method (OEM).

The name has been changed according to suggestion.

PAGE 2:

- *Unfortunately, instruments available nowadays are not able to provide sufficient resolution to describe short time scale processes such as convection, cloud formation or boundary layer turbulence.* if you are going to make a blanket claim like this you must defend it. I suggest you remove it.

Modified in the manuscript.

- Raman Lidar (RL) be specific here, vibrational or rotational?

There is no need to specify because it is the two of them: the water vapor Raman Lidar is based on VRR, and the temperature profiling is based on a RR scattering. (R = Rotational, V= Vibrational)

- cannot provide information about water vapour, the do provide information!

In the text it is written that the lidar cannot provide information **above and within optically thick clouds**. That holds true. Even though the instrument could provide information above a thin enough cloud.

More specifically, a ground-based Raman lidar with a very powerful laser, as the one considered in this study, can provide information on atmospheric humidity within optically thick clouds. Nevertheless, it holds only for 100-200 m till the laser beam gets completely extinguished. This typically happens for an UV optical thickness of the cloud up to 1-2.

- *which drastically reduces the quality of the data.* drastic is too strong a word, and perhaps it is more the "quantity" that suffers as opposed to quality.

Corrected in the manuscript.

- *the information of the lowest layers in the atmosphere cannot be used*, how low, overlap is highly system depend and many different solutions exist, you have to qualify this is for some lidars, not others.

Corrected in the manuscript.

- *there are no lidar data*. with your system!

It is an unavoidable lidar feature. Because the receiver and the transmitter systems have a slightly bistatic configuration, there is a region where the field of view of the telescope does not superimpose with the laser beam. In this region, the instrument is not capable of providing information. It is well known that, by performing ratios between different channels, one could partially get rid of some problems related to the overlapping function (OVF). For example, in the BASIL system, the water vapor initial OVF is as large as almost 0.7 km. Nevertheless it can be reduced up to around 180 meters, which is very good, but not enough: we still need to get a good estimate of the water vapor from ground to 180 meters, region where the water vapor variation is especially strong. This is the region where MWR can provide good information content, and this is one of our strongest arguments to perform the synergy.

PAGE 3

- *A method to combine RL and MWR was already proposed by Han et al. (1997), where the authors developed a two-stage algorithm to derive water vapor atmospheric profiles. In the first stage, a Kalman filtering algorithm was applied using surface in situ and RL measurements. In the second stage, a statistical inversion technique was applied to combine the Kalman retrieval with the integrated water vapor of a two-channel MWR and climatological data. Their method showed that the synergy of these two sensors compensate for the individual sensor's drawbacks. A continuation of this work was carried out by Schneebeli (2009) where, still following the Kalman filter two-stage configuration, the products were extended to also temperature profiles.*

you need to include and discuss:

Leblanc, T. (2008), Accuracy of Raman lidar water vapor calibration and its applicability to long-term measurements, Appl. Opt., 47(30), 5592–5603.

which compares with radiometers and sondes as well as:

Venable, D. D., D. N. Whiteman, M. N. Calhoun, A. O. Dirisu, R. M. Connell, and E. Landulfo (2011), Lamp mapping technique for independent determination of the water vapor mixing ratio calibration factor for a Raman lidar system, Appl. Opt., 50(23), 4622–4632, doi:10.1364/AO.50.004622.

who do a full internal calibration of their lidar.

Similar references are already included and discussed in the lidar section (2.1).

- *supersites* don't use italics for this

Corrected in manuscript.

PAGE 4

- *They can be traded-off to improve measurement precision, with random error in the measurements being inversely proportional to the square root of both vertical and temporal resolutions. The height-time resolution product can be varied to improve SNR...,*
  but it is probably not necessary to state that if you don't want to, but if so word it better.

Rephrased in the manuscript.

- *presence of a blind region in the lower altitudes, vertical profiles of* do you mean because of the geometry of the large telescope-transmitter system the wv...

Corrected in the manuscript.

- *drawback* limitation

Corrected in the manuscript.

- *might* ? do they or don't they?

They might, it is situation dependent.

- *profiles* profiles of what?

Clarification in text included.

- *approximately 4 km away from the instrument.* you've mentioned that, but what is relevant is where the sonde is at the heights the calibration is made, no necessarily where you launch from

Yes, there is a drift inherent to the sonde flight. The lidar calibration is always performed trying to get the best reference, but the exact distance to the lidar is different in every calibrated profile. That is the reason why we refer typically to the launch distance, which is constant.

PAGE 5

- *to not exceed 5%. Considering a vertical and temporal resolution of 150 m and 5 min, respectively, the statistical error affecting water vapor mixing ratio measurements*

for night-time operation is typically smaller than 2% up to 3 km and smaller than 20% up to 9 km, while for daytime operation is typically smaller than 40% up to 3 km and smaller than 100% up to 4.5 km. Additionally, the statistical error affecting temperature measurements for night-time operation is typically smaller than 0.4 K up to 3 km and smaller than 1 K up to 6.5 km, while for daytime operation is typically smaller than 0.5 K up to 3 km and smaller than 1 K up to 4.5 km. In addition to the statistical error, a small systematic error (bias) may affect the water vapor and temperature measurements. For example, for water vapor measurements, besides a bias associated with the estimate of the calibration coefficient itself (radiosonde biases, different air masses being sensed by the radiosonde and the lidar), an additional bias (<1%) may be associated with the use of narrowband filters, the temperature dependence of H2O and N2 Raman scattering and the drifts of the filters position associated with thermal drifts (Whiteman, 2003). Still an additional 1% may be associated with the determination of the differential transmission term at the water vapor and molecular nitrogen Raman wavelengths (Whiteman, 2003). lots of information here! How about putting it in a Table as well so it is easier for the reader to grasp.

Rephrased in the manuscript.

- *two frequency bands: K and V bands* in the K and V frequency bands.

Corrected in text.

PAGE6

- *low amount of vertically independent information (i.e. 2 pieces of information per profile for water vapor, typically 3-4 for temperature)* ... corresponding to vertical resolutions of ???

In general, the degrees of freedom of signal per profile do not depend on the vertical discretization. Regarding vertical resolution of the retrieved humidity profile, a figure showing it for an exemplary profile has been added to the other document.

- *apriori* a priori is 2 words

Corrected in the manuscript.

- *the moderately non-linear nature of our problem, the iterative equation applied to find the best atmospheric state estimate is:* Please define what you mean by moderately non-linear, that is not a standard term, put Rodgers references at end of sentence.

Corrected in the manuscript.

- *Where* typo should be where and flush to margin

Corrected in manuscript.

- *y is the observation vector* The observation vector, *y*, contains...

Corrected in manuscript.

- *in our case, coming from radiosondes* specifically which quantities from the radiosondes?

Detailed description is included in section 3.2. “Apriori”.

- *F (xi) is the forward model applied to the state vector xi, whose output lies on the observation space* in general the forward model can depend on both retrieved and model parameters.

The text has been changed to explicitly state the dependency of the forward model to the model parameters.

- *which can be understood as the variation on the observation vector when a perturbation is performed on the state vector (eq. (2)).* ? your equation says the Jacobian is the variation of the FM wrt the retrieval parameter?

The equation defines the K matrix as the derivative of the observation vector wrt the state vector. The manuscript has been changed to clarify this definition.

PAGE 7

- *The iterative equation described in (1) finds the most optimal atmospheric state xop. This state is reached if the convergence criterium is fulfilled (Rodgers, 2000): d2i = (yi+1 − yi)T (S ε (KSaKT + S ε)S ε)−1(yi+1 − yi) ≪ m (3) where m is the number of elements in the observation vector and much smaller refers to at least one order of magnitude smaller.* You seem to be missing the key equation to OEM, that for the cost. Cost is what is relevant, you can converge but still have a high cost, meaning your residuals are not white noise.
  Please include the cost function here, as well as telling us the cost when you present retrievals.

The authors agree that convergence doesn't mean having a small cost. Please check the plots on the other document were this issue is addressed.

- *error estimation* be specific here, error due to what?

The error S_{op} corresponds to the a posteriori covariance matrix for the solution with expected value x_{op}. It is derived from measurement uncertainties, a priori uncertainty and Jacobian (see equation (4)).

- *t is the temperature* I strongly urge you to follow the convention of using *T* for temperature and *t* for time, otherwise your paper will be very difficult to read

Corrected in the manuscript.

PAGE 8

- *The vector y is composed of the TBs from the MWR and the mixing ratio and/or temperature from the RL.* don't use bold unless these are matrices. What is a TB? Did you define this earlier?

Bold corrected in the manuscript.
TB is previously defined at line 186 as Brightness Temperature.

- *Therefore, before the analysis, a running average is performed on the data (we choose 300 m window size) previously to the analysis.* I'm not sure this is correct, but I will leave it up to the referees. You have a measurement and data grid. You should make you retrieval grid 300 m and your data grid 30 m, or make the data grid 300 m and then some other choice for the retrieval grid.

I am concerned you are running a boxcar smoothing on the data which essentially is introducing additional correlations and affecting the \epsilon term in the retrieval. You

are discussing the measurement vector, y, which is related to the FM by the noise, which I am concerned you are playing with here, e.g.:

y = F(x,b) + \epsilon

See answer on reply 1 of the other document.

PAGE 10

- This definition implies no correlation between measurements 275 in different heights. then it is composed of the variances, not the covariances

Corrected in the manuscript.

- *Allows to work with one* works with a single

Corrected in the manuscript.

PAGE 11

- *In the following, the results for one example profile are presented (Fig. 2).* for this figure to be meaningful I would need to see the following:

- the Jacobians
- the averaging kernels
- the vertical resolution with height
- the residuals (since we don't know the cost we need to see the residuals are white and don't have biases)

Please, see answer number 2 in the other document.

- *At first, we introduce in the OES only the portion of profile where RL data is valid (i.e. from 180 m to 2.5 km, ~ 44layers), not taking into account the MWR. The result of the algorithm is a complete profile from ground to 10 km. In the region with lidar availability, the result will tend to the 305 portion of lidar profile, since the error associated to this measurements is very small (on the order of 0.5g/m3). In the regions where no lidar data can be defined, the profile will be completed with the information provided by the apriori profile, which is the only information available.* no one doubts you, but we must see the averaging kernels to know where this is coming from.

See answer number 2 of the other document.

PAGE 12

- *This might be explained because the sonde has been launched under different local conditions: while the instruments site is located inside the research center, the RS is launched in an open field area.* It could cause slight differences in the retrieval close to the ground, but should not be a problem in the free troposphere.

  or it could be caused by a height variation in the lidar instrument function which occurs from calibrating at a limited range of heights.
  Please comment on this in the text.

There is no height variation in the lidar instrument function. As mentioned already in the manuscript, the overlap function is the same for the two channels considered for the water vapour Raman lidar measurements (water vapour and molecular nitrogen Raman channels). They consequently cancel when calculating the two channels ratio. We have carefully

checked in the past for the presence of a differential overlap between the two channels, but we could not find any. The optical layout of our receiving system is very compact. This feature strongly reduces the risk in this direction. This has been carefully verified in the past against collocated radiosonde data. So calibration, even if performed at a limited range of heights, which is not the case, cannot determine a height variation in the lidar instrument function.
Additionally, the figure under discussion clearly reveals that the water vapour profile from the lidar is in very good agreement with the other collocated instrument, i.e. the MWR, while the disagreement is only with the only instrument which is not collocated, i.e. the radiosonde. This testifies that the disagreement between the radiosonde and the lidar/MWR is much more likely due to the different air masses sampled by the instruments, within this height interval.

- *Out of all the clear sky profiles, the convergence of the OES is found in 82% of the cases, that is, 687 profiles. In the rest of the cases, the convergence is not found because the algorithm cannot find a profile which is simultaneously consistent with the measurements of the two instruments and the apriori, within their uncertainties.* does this result make you concerned about why this occurs, and whether your "good" results are in fact unique?

  Perhaps inspection of the averaging kernels will reveal why this method is "marginally stable", in the sense that it should be more robust?

The number 82% has been corrected: there was an error in the count of the total number of profiles. The real percentage is a convergence of 95% instead of 82%. It has been corrected in the manuscript.
There is a remaining <5% of non-converging profiles. As mentioned on the manuscript (lines 342 to 345), the convergence is typically not possible because no agreement between the two instruments is found (inside the instruments uncertainty). This is typically due to biases in the lidar profile.

PAGE 13

- *deviation smaller than 1.2kg/m2. These values lie inside the GPS uncertainty of 1−2kg/m2 (Gendt et al., 2004) and the MWR product of ∼ 0.5 − 1kg/m2 (Steinke et al., 2014). This result gives us confidence that the developed OES method delivers reliable water vapor profiles.* is the comparison of the IWV with your retrieval better or worse than the comparison of GPS with just the HATPRO, in other words does your retrieval increase or decrease the IWV from just the radiometer?

The mean difference and deviation are described in Table 1. Since the average difference between our method and the MWR is positive, one can read that the retrieval increases slightly the IWV from just the MWR.

PAGE 14

- *Figure 5. Mean and standard deviation of the difference between the 18 clear sky radiosondes: MWR (in green), RL (in red) and the combination of both (blue). The dashed horizontal lines enclose the region where the lidar data is used.* this curve doesn't mean much in these units, what we want to know if what is the percent difference.

We include this graph in absolute difference for two essential reasons:

- first of all, to compare this result to the theoretical error, which is also expressed in g/m3. This comparison is crucial.

- Second, it is needed to justify the factor of 4 in our "Increase of the RL error" section.

PAGE 15

- *Region a) from ground to 180 m* , data is available from ? and ? do this as well for Region C

Corrected in manuscript.

- *In region (c) all the three values for the different retrievals are similar. The only-MWR seems to perform best when comparing to the RS, because both its bias and stv are the smallest. The only- RL case presents the largest bias and stv because in this region only information from the apriori is provided. The combination of the two sensors presents intermediate values.* since it is the relative biases/SD that are relevant it is impossible for me to access this without knowing what they are, particularly in regions b & c.

We believe that the comparison between the RS and the three instruments is already properly addressed by showing the absolute difference. The justification to show the absolute difference is explained above.

PAGE 16

- *Unfortunately, a set of only 18 radiosondes is not enough to asses the benefits of the synergy.* please justify, reference or delete this assertion.

From a rigorously statistical point of view, a number of 18 samples is not enough to be considered statistically significant.

- *Launch distance of 4 km to the site, drifting of the balloon, dry bias, etc.* actually aren't these not the hard to assess? For instance, isn't there a barometer and maybe a flux tower near the lidar? You can compare that to the place the sonde was launch and see how similar or different it is. As far as the drift of the balloons, the balloons are tracked so you can easily tell if they are flying over you or not. As far as biases there are many references on this, please quote the relevant papers in the literature.

  I believe there is plenty of real information to be used and then compared to a theoretical treatment.

See answer number 4 of the other document.

- *As already mentioned in section 3, the algorithm provides an estimation of the error for the retrievals, see eq. (4). This theoretical error is computed for every profile and for the three different cases: using only-RL, only MWR and the RL+MWR combination.* as equation 4 shows we need to see the averaging kernels to assess whether what is being shown in Fig 6 makes sense or not.

Replied on the other document, answer number 2.

PAGE 17

- *lidar OVF in this region.* the standard water vapour analysis divides the water vapour counts by the nitrogen counts. This the overlap function does NOT appear in the retrieved water vapour for an achromatic system.

  If you believe the *differential* overlap is an issue you need to discuss this in the instrument section (have any studies of this been made for this system).

Also, you have not discussed uncertainties in your knowledge of the extinction in the PBL, whose affect can be much larger than the differential overlap. Please discuss this as well.

See answer number 5 of the other document.

PAGE 19

- *As explained in section 3.3, only Poisson noise was taken into account but there can be other possible sources of uncertainty.* please state what there are and what there magnitude is

These sources of uncertainty are detailed explained in section 2.1.

- *we artificially incremented the RL error by a factor of 4 to study the sensitivity of the retrieved profile with respect to the RL measurement uncertainty.* I don't think this is reasonable, why not 3.6674 or 4.12839? You should be able to estimate a "theoretical" systematic lidar error budget and use that.

The explanation is found in lines 447-450. In addition, a detailed explanation is included in the other document answer number 6.

PAGE 20

- *Degrees of freedom for signal comparison for absolute humidity. Average over 636 profiles. The atmosphere is separated in three regions according to lidar availability. The DOF are presented for three cases: only RL, only MWR and the combination of both instruments. In the upper part, no increment on the RL error has been considered. In the bottom part, the RL uncertainty has been multiplied by a factor of four.* these numbers seem odd and one might question whether they are correct. Perhaps showing the averaging kernels here would clear up why some of the numbers appear unusual.

The authors do not believe the numbers included in table 2 are odd. We hope that, by including the cumulative DOF plot, the interpretation of the DOF is easier. For clarity, we also included plots for the averaging kernels in the other document.

PAGE 21

- *In the case of the temperature, the lidar profile for this specific case study is much more affected by the OVF than in the water vapor profile, and so there are no valid temperature lidar measurements under 2.5 km.* perhaps you can refer to another paper or show a curve of the typical overlap function of this lidar, why is the differential overlap so large near 2.5 km, especially if you have measurements from 2 channels with smaller telescope to correct it with.

  Please explain in the text why this is so.

Temperature measurements by rotational Raman lidar can be obtained from the power ratio of the high-to-low quantum number rotational Raman signals. This ratio R(z) is related to atmospheric temperature through the following analytical relationship:

$$R(z)=\exp(a/T+b)$$

with a and b being two calibration constants. Significant changes in atmospheric temperature can result from small changes of R(z), because of the non linear relationship between R(z) and T. In this respect, small differences in the overlap function for the two channels can result in non-negligible systematic effects for temperature, especially in the

situation of a poor characterization of the overlap function. We prefer to stay on the safe side and neglect the data with high noise values.

PAGE 22

- *Table 3. Degrees of freedom for temperature retrieval, separated in three regions in the atmosphere. Lidar data is only present in region (b). The DOF are presented for the cases were only-RL is used, only-MWR and for the combination of the both instruments* as before we need to see Jacobians, residuals, averaging kernels, vertical resolution

See answer number XXX of the other document.

- *The degrees of freedom for the temperature profiles are also presented in table 3. The independent pieces of information are improved in the lower part of the atmosphere when introducing MWR information. The combination RL-MWR presents the highest information content, increasing the number of DOF in more than one, with respect to the only-RL case, and in ∼ 6, with respect to the only-MWR profile.* you water vapour method was unstable, in the sense it didn't work for quite a few cases. I would like to see a few other nights of temperatures to get an idea of whether this retrieval was tuned for a single night but doesn't work in general.

See answer number 7 of the other document.

- *T and Q a* ? what are T and Q, the former t and q?

Corrected in manuscript.

PAGE 23

- *Absolute humidity, temperature and relative humidity from RS (black), profiles retrieved separately using MWR+RL (blue), the simultaneous T-Q retrieval using MWR+RL(red) and the simultaneous T-Q retrieval without RL temperature (yellow). Horizontal bars represent the error associated to the resulting profiles. The horizontal grey dashed lines enclose the areea where lidar data was available. Numbers represent the averaged difference to the RH of the RS for each case in percentage [%].* same story, we need the jacobians, averaging kernels, residuals to evaluate this

See answer number 2 of the other document.

PAGE 24

- sky measurements?

Corrected on the manuscript.

- *The improvements of the synergy have been analysed in terms of several parameters, like the reduction of the theoretical error or the increase of DOF, showing strong advantages with respect to the two instruments working separately. For example, when applying the combined retrieval to the complete HOPE period, the absolute humidity error can be reduced by 59.8% and 37.9% on average, with respect to the retrieval using only MWR data or only RL, respectively. Results for a case study temperature profile show that the error is improved in a 47.1% and 24.6% with respect to the only-MWR and only-RL profiles, respectively. The synergy present*

its strongest advantages in the regions where RL data is not available, whereas in the regions where both instruments are available, RL dominates the retrieval. I am deeply concerned these results are based on "theoretical" errors while the study is using real measurements which should have real uncertainties that can be estimated (and then compared to the theoretical).

As already discussed, the availability of reference profiles to compare the results with is reduced. We only have 18 possible profiles of RS to study, which is a very reduced number to provide a good long term statistics.

- It would also be nice to see some scientific result from all this techniques, in the sense of some new or improved understanding of the atmosphere.

This is a very good suggestion. Indeed, there is an ongoing effort to perform the suggested study. Nevertheless, the aim of the paper is to present the method and prove its applicability. The authors understand that the application of the algorithm for atmospheric processes, fills other complementary study.

Barrera-Verdejo (mbarrera@smail.uni-koeln.de)
Manuscript prepared for J. Name
with version 2014/09/16 7.15 Copernicus papers of the LaTeX class copernicus.cls.
Date: 11 April 2015

Ground Based Lidar and Microwave Radiometry Synergy for High Vertically Resolved Thermodynamic Profiling.

María Barrera-Verdejo[1], Susanne Crewell[1], Ulrich Löhnert[1], Emiliano Orlandi[1], and Paolo Di Girolamo[2]

[1]Universität zu Köln
[2]Scuola di Ingegneria, Universita degli Studi della Basilicata

Correspondence to: María

Abstract. Continuous monitoring of atmospheric humidity and temperature profiles is important for many applications, e.g. assessment of atmospheric stability and cloud formation. While lidar measurements can provide high vertical resolution albeit with limited coverage, microwave radiometers receive information throughout the troposphere though their vertical resolution is poor. In order to overcome these specific limitations the synergy of a Microwave Radiometer (MWR) and a Raman Lidar (RL) system is presented in this work. The retrieval algorithm that combines these two instruments is an Optimal Estimation Method (OEM) that allows for a uncertainty analysis of the retrieved profiles. The OEM combines measurements and a priori information taking the uncertainty of both into account. The measurement vector consists of a set of MWR brightness temperatures and RL water vapor profiles. The method is applied for a two month field campaign around Jülich, Germany for clear sky periods. Different experiments are performed to analyse the improvements achieved via the synergy compared to the individual retrievals. When applying the combined retrieval, on average the theoretically determined absolute humidity error can be reduced by 59.8% (37.9%) with respect to the retrieval using only-MWR (only-RL) data. The analysis in terms of degrees of freedom for signal reveals that most information is gained above the usable lidar range. The retrieved profiles are further evaluated using radiosounding and GPS water vapor measurements. Within a single case study we also explore the potential of the OEM for deriving the relative humidity profile, which is especially interesting to study cloud formation in the vicinity of cloud edges. To do so temperature information is added both from RL and MWR. For temperature, it is shown that the error is reduced by 47.1% (24.6%) with respect to the only-MWR (only-RL) profile. Due to the use of MWR brightness temperatures at multiple elevation angles, the MWR provides significant information be-

low the lidar overlap region as shown by the degrees of freedom for signal. Therefore it might be sufficient to combine RL water vapor with multi-angle, multi-wavelength MWR for the retrieval of relative humidity, however, long-term studies are necessary in the future. In general, the benefit of the sensor combination is especially strong in regions where Raman Lidar data is not available (i.e. overlap region, poor signal to noise ratio), whereas if both instruments are available, RL dominates the retrieval.

1 Introduction

Humidity and temperature are essential variables for the description of any meteorological process. Highly resolved, accurate and continuous measurement of these parameters, in particular water vapor, are required for a deeper understanding of many atmospheric phenomena (Stevens and Bony, 2013). Unfortunately, instruments available nowadays are not able to capture humidity and temperature with sufficient spatial and temporal resolution to describe short time scale processes such as convection, cloud formation or boundary layer turbulence.

Nevertheless, in order to overcome the specific limitation of a specific instrument, the scientific community started merging different data from several instruments in the last years. Some examples of these are Stankov (1998) or Löhnert et al. (2001), where information from diferent sources is combined. In the present paper, the synergy between ground based Raman Lidar (RL) and Microwave Radiometer (MWR) is described. Both instruments present some advantages and disadvantages and, by bringing them together in an optimal and new retrieval algorithm, it is possible to overcome some of the dissadvantages in the single devices and enhance their benefits.

The Raman lidar systems provide highly resolved measurements of atmospheric humidity profiles. For this reason, Raman lidars have become a strong tool for active ground based observations in the last years. However, the RL technique presents important weaknesses which prevent it from effective operational application. For example, ground based RL cannot provide information above and within optically thick clouds, as the radiation emitted by the lidar gets attenuated once the laser beam reaches the liquid layers within the cloud. Moreover, day time measurements are affected by background solar radiation, which strongly reduces the quality of the data. The continuous and effective detection of Raman signals, which are especially weak, requires robust and stable alignment of the receiving system. Daytime operation requires the use of powerful lasers whose continous operation is technically demanding. Additionally, RL needs to be calibrated. This calibration is usually performed based on the use of radiosounding data, which presents some caveats. First, the balloon might measure a different air volume due to its drift. Second, it implies a high both human and instrument cost. In addition, when measuring with the lidar, the information of the lowest layers in the atmosphere typically cannot be used, due to the presence of a blind region associated with the overlap function (OVF) of the RL.

The MWR allows continuous passive data acquisition and it is a robust operational instrument (Rose et al., 2005), measuring unattended in a 24/7 mode. In contrast to RL, the instrument offers a much more limited vertical resolution of the retrieved atmospheric profiles, especially in higher layers of the atmosphere (i.e. above an altitude of 1km) (Löhnert et al., 2007), but performs best for measurements close to the ground, where there are no lidar data. MWR also provides accurate integrated quantities such as Integrated Water Vapor (IWV) or Liquid Water Path (LWP). The calibration of this instrument is easily performed with internal references with known temperature (hot load-cold load) or by observing the atmosphere under different elevation angles (i.e. sky tipping) (Maschwitz et al., 2013). Another advantage of the MWR is the capability of measuring in almost all weather conditions (also cloudy cases) except for rainy scenarios, where the received signal must be discarded in most of the cases.

A method to combine RL and MWR was already proposed by Han et al. (1997), where the authors developed a two-stage algorithm to derive water vapor atmospheric profiles. In the first stage, a Kalman filtering algorithm was applied using surface in situ and RL measurements. In the second stage, a statistical inversion technique was applied to combine the Kalman retrieval with the integrated water vapor of a two-channel MWR and climatological data. Their method showed that the synergy of these two sensors compensate for the individual sensor's drawbacks. A continuation of this work was carried out by Schneebeli (2009) where, still following the Kalman filter two-stage configuration, the products were extended to also temperature profiles.

The method described in this document is a new approach based on an Optimal Estimation Method (OEM), an iterative optimal and physically consistent method that allows uncertainty assessment and provides the most probable estimated atmospheric state together with its uncertainty description. The aim of this study is to combine the information provided by the two instruments in an OEM to retrieve atmospheric parameters. The method was applied to the data collected during HOPE ($HD(CP)^2$ Observational Prototype Experiment), focusing on clear sky cases. Results for absolute humidity (AH), temperature (T) and relative humidity (RH) profiles are shown. A detailed description of the method is presented in section 3. Section 4 describes the results when applying the method to retrieve absolute humidity profiles: both for a single case study and the complete two months period of HOPE. In addition, an example of temperature retrieval will be presented in section 5. Moreover the algorithm is used to simultaneously retrieve absoulte humidity and temperature profiles, which leads to the calculation of the relative humidity profile (see section 6). Finally, section 7 summarizes the results and provides an outlook.

2 Observations: HOPE

In this study we make use of the data collected during HOPE, which was a major field campaign in Nordrhein-Westfalen, Germany, from April to June 2013. The main goal of the campaign was

to provide a complete picture of the clouds lifetime and evolution. During the measurement period, three supersites were operating, distributed in the surroundings of Forschungszentrum Jülich. Each supersite was composed of a rich variety of remote sensing instruments, coordinated with different scanning strategies that allow the 3D study of clouds.

At the supersite of JOYCE (Jülich ObservatorY for Cloud Evolution) (Löhnert et al., 2014), measurements by the University of Basilicata Raman Lidar system (BASIL) and a MWR were carried out. Also, auxiliary data from other instruments is available and, in addition, a large set of radiosondes (RS). The RS set is composed of more than 200 sondes, launched only 4 km away from JOYCE, typically at least twice a day.

2.1 BASIL

The Raman Lidar system BASIL (Girolamo et al., 2009),(Di Girolamo et al., 2012) is an active instrument based on the detection of the elastic and Raman backscattered radiation from atmospheric constituents. BASIL includes a Nd:YAG laser emitting pulses at its fundamental wavelength, its second and third harmonics: 1064, 532 and 355 nm, respectively. Raman scattering is stimulated by the 355 nm wavelength, a frequency of 20 Hz, with an average power emitted at this wavelength of 10 W. The receiver is built around a larger telescope in Newtonian configuration (45 cm diameter primary mirror) and two smaller telescopes (5 mm diameter lenses). The larger telescope is primarily dedicated to the collection of the Raman signals, i.e. the water vapor and molecular nitrogen roto-vibrational Raman signals, at 407.5 and 386.7 nm, respectively, which are used to estimate the water vapor mixing ratio profiles; and the molecular nitrogen and oxygen pure-rotational Raman signals, at 354.3 and 352.9 nm, used to estimate the atmospheric temperature profiles.

Signal selection is performed by means of narrowband interference filters, whose specifications were reported in Di Girolamo et al. (2004) and Girolamo et al. (2009). Sampling of the Raman signals is perfomed by means of transient recorders with double signal acquisition mode (i.e. both analog, A/D conversion and digital, photon counting). Depending on the application, water vapor mixing ratio and temperature profiles can be derived with different vertical and temporal resolutions. These two parameters can be traded-off to improve measurement precision. For the purposes of this study, the lidar products are characterized by a vertical resolution of 30 m and a temporal resolution of 5 minutes. Because of the geometry of the telescope-transmiter system, there is a blind region in the lower altitudes. Due to that, vertical profiles of water vapor mixing ratio typically start at 150-180 m; and temperature profiles at around 300 m. This limitation is caused by OVF problems and is due to a non sufficient overlap between the lidar emitter and receiver systems. Nevertheless, temperature profiles might present problems with overlaping function until $\sim 1.5km$. Temperature and humidity profiles extend vertically up to different altitudes during daytime and night-time depending on when the signal gets completely extinguished. For water vapor this typically takes place around 4 km

during daytime and around 10 km during the night, while for temperature it typically takes place around 6 km during daytime and up to 20 km during the night.

During HOPE, BASIL has been calibrated based on the comparison with the radiosondes launched approximately 4 km away from the instrument. A mean calibration coefficient was estimated comparing BASIL and radiosonde data. This data is compared in an altitude region with an extent of 1 km above the boundary layer to minimize the air mass differences related to the distance between the lidar and the radiosonde. The standard deviation of the mean calibration coefficient from the single values does not exceed 5%. We have considered a vertical and temporal resolution of 150 m and 5 min respectively. With this values, the statistical error affecting water vapor mixing ratio measurements for night-time operation is typically smaller than 2% up to 3 km and smaller than 20% up to 9 km. While for daytime operation is typically smaller than 40% up to 3 km and smaller than 100% up to 4.5 km. Additionally, the statistical error affecting temperature measurements for night-time operation is typically smaller than 0.4 K up to 3 km and smaller than 1 K up to 6.5 km. While for daytime operation is typically smaller than 0.5 K up to 3 km and smaller than 1 K up to 4.5 km.

In addition to the statistical error, other small systematic error sources might affect the water vapor and temperature measurements. For example, for water vapor measurements, besides a bias associated with the estimate of the calibration coefficient, an additional bias (<1%) might be considered. This percentage is associated with the use of narrowband filters, the temperature dependence of H2O and N2 Raman scattering and the thermal drifts of the filters (Whiteman, 2003). Still an additonal 1% might be associated with the determination of the differential transmission term at the water vapor and molecular nitrogen Raman wavelengths (Whiteman, 2003). This sources of error, in principle negligible, are not taken into account for the calculations in our algorithm.

The operation of BASIL has not been continuous during HOPE, the instument has collected a total of 430 hours of measurements distributed over 44 days, which represents the 30% of the whole HOPE period.

2.2 MWR

The microwave radiometer profiler HATPRO (Rose et al., 2005) was manufactured by Radiometer Physics GmbH, Germany (RPG) as a network-suitable microwave radiometer with very accurate retrievals of Liquid Water Path (LWP) and Integrated Water Vapor (IWV) at high temporal resolution (1 s) (Löhnert and Maier, 2012). It is a passive MWR that measures radiation in the atmosphere in two frequency bands in the K and V bands (Rose et al., 2005). The seven channels of the K band contain information about the vertical profile of humidity through the pressure broadening of the optically thin 22.235-GHz H_2O line and contain also information for determining liquid water path. The seven channels of V-band contain information on the vertical profile of temperature resulting from the homogeneous mixing of O_2 throughout the atmosphere (Löhnert et al., 2009).

The absolute calibration of the instrument is performed taking a cold and a hot load as references, which are assumed to be ideal black bodies. The cold body is a liquid-nitrogen-cooled load that is attached externally to the radiometer box during maintenance, which can be considered as a black body at the LN_2 boiling temperature of aproximately 77 K. This standard, together with an internal ambient black body load, is used for the absolute calibration procedure (Maschwitz et al., 2013). In addition, a calibration by tip-curve observations is performed, whereby the instrument collects observations for K-band channels at different elevation angles (Turner et al., 2007). The reliability of sky tipping calibrations will strongly depend on how good the assumption of an horizontally stratified atmosphere is. Further details on the calibration procedures of the instrument can be found in Maschwitz et al. (2013).

For the water vapor study only zenith measurements have been used since non-zenith measurements do not improve the retrieval of vertical humidity profiles (Löhnert et al., 2009). But for the temperature retrieval, angular information can be used to improve the accuracy and vertical resolution of the retrieved profile (Crewell and Lohnert, 2007).

The temporal resolution of this instrument is higher than for the RL: it is able to provide one measurement every second, so a temporal adaptation to the lidar time resolution is performed, averaging MWR measurements in five minutes intervals. A major drawback of MWR retrievals is the low amount of vertically independent information (i.e. 2 pieces of information per profile for water vapor, typically 3-4 for temperature) (Löhnert et al., 2007).

3 Method

3.1 Optimal Estimation Method

An Optimal Estimation Method allows to estimate the state of the atmosphere and its associated uncertainty. Using this scheme requires a set of measurements (with their error specification), a forward model for calculating the atmospheric state from the measurements, and some*a priori* information. In the following, a short description of the scheme is presented. Deeper details can be found in Rodgers (2000).

Given the *moderately non-linear* nature (Rodgers, 2000) of our problem, the iterative equation applied to find the best atmospheric state estimate is (Rodgers, 2000):

$$x_{i+1} = x_a + (S_a K_i^T (K_i S_a K_i^T + S_\epsilon)^{-1} [y - F(x_i) + K_i(x_i - x_a)]) \quad (1)$$

where x_i is a vector containing the atmospheric state at the iteration i, that is: the profiles of temperature and/or humidity. The observation vector y contains the brightness temperatures (TB) from the MWR and the mixing ratio or temperature from the lidar. The term x_a represents the *a priori* information of the atmosphere, in our case, coming from radiosondes. S_a and S_ϵ are the

covariance matrices of the prior and observation uncertainties, respectively. $F(x_i, b)$ is the forward model applied to the state vector x_i, and depending on the model parameters b. For simplicity, it will be referred as $F(x_i)$. The forward model output lies on the observation space. The term K represents the Jacobian, which can be understood as the variation on the observation vector when a perturbation is performed on the atmospheric state vector (eq. (2)):

$$K_i = \frac{\partial F(x_i)}{\partial x_i} \tag{2}$$

The iterative equation described in (1) finds the most optimal atmospheric state x_{op}. This state is reached if the convergence criterium is fulfilled (Rodgers, 2000):

$$d_i^2 = (y_{i+1} - y_i)^T (S_\epsilon (K S_a K^T + S_\epsilon) S_\epsilon)^{-1} (y_{i+1} - y_i) \ll m \tag{3}$$

where m is the number of elements in the observation vector and much smaller refers to *at least one order of magnitude smaller*. An error estimation of the solution S_{op} is calculated via:

$$S_{op} = S_a - S_a K^T (S_\epsilon + K S_a K^T)^{-1} K S_a \tag{4}$$

where K is the Jacobian calculated in the last iteration. It is also posible to estimate the information content of the result. The degrees of freedom (DOF) of a profile represent the amount of independent pieces of information in the signal. They can be calculated as the trace of the matrix in the following equation (5) (Rodgers, 2000):

$$A_{ker} = S_a K^T (S_\epsilon + K S_a K^T)^{-1} K \tag{5}$$

where A_{ker} is the averaging kernel. This matrix is very important to describe the information content, as it describes the subspace of *state space* in which the retrieval must lie (Rodgers, 2000).

3.2 *A priori* calculation: x_a and S_a

The a priori information is calculated from the set of radiosondes launched during HOPE. A total of 217 sondes have been considered as valid. Generally, at least two of them are available for every day of the campaign, typically one around noon and the other at midnight. From these data, average profiles of temperature (T) and humidity (q) have been calculated to represent the a priori knowledge, together with their standard deviation. These profiles represent x_a in the algorithm described by eq. (1).

For the same set of radiosondes, the correlation and covariance (S_a) matrices are calculated according to(Wilks, 2006):

$$S_{a,(T,q)} = \begin{pmatrix} cov(T,q) & cov(q,q) \\ cov(T,T) & cov(q,T) \end{pmatrix} \tag{6}$$

where q is the absolute humidity and T is the temperature defined as a function of the altitude:

$$\begin{aligned} q &= [q_1, q_2, ...q_k] \\ T &= [T_1, T_2, ...T_k] \end{aligned} \tag{7}$$

and k is the total number of altitudes in the retrieval vertical grid. Both covariance (cov) and correlation (corr) matrices have been calculated as in equation (8). The covariance matrix is calculated because it is needed in the algorithm as input (S_a), the correlation matrix because it better illustrates the relations between water vapor and temperature in the atmosphere.

$$corr_{ab} = \frac{cov(a,b)}{s_a s_b} = \frac{\frac{1}{n-1}\sum_{i=1}^{n}[(a_i - \bar{a})(b_i - \bar{b})]}{\left[\frac{1}{n-1}\sum_{i=1}^{n}(a_i - \bar{a})^2\right]^{\frac{1}{2}}\left[\frac{1}{n-1}\sum_{i=1}^{n}(b_i - \bar{b})^2\right]^{\frac{1}{2}}} \tag{8}$$

where i gOEM over each radiosonde, with a total of $n = 217$. a and b represent both absolute humidity and/or temperature profiles. The parameters $\bar{a}$ and $\bar{b}$ are the averaged vertical profiles for temperature and/or absolute humidity.

The correlation matrix is presented in Figure 1. It shows how the two variables (q,T) are correlated as a function of the altitude, from ground to 10 km, and is composed of the four submatrices: corr(T,T), corr(q,q), corr(q,T) and corr(T,q).

The temperature corr(T,T) clearly shows the tropopause at altitudes $> 9km$. The corr(T,T) values are higher than the water vapor corr(q,q) values, which show a much higher variability. The values for corr(q,q) are strongest close to the main diagonal, but decrease quickly for off diagonal terms, whereas the corr(T,T) is stronger in the off diagonal terms. In the lowest 1-2 km there is a higher correlation in all cases, because of the well mixed conditions in the boundary layer. The results are similar to previous studies Ebell et al. (2013).

In this study, the submatrices $S_{a,(q,q)} = cov(q,q)$ and $S_{a,(T,T)} = cov(T,T)$ will be used in Sections 4 and 5, respectively, when only absolute humidity or temperature are retrieved separately. The complete matrix $S_{a,(T,q)}$will be needed for the simultaneous retrieval of the two atmospheric states, in Section 6.

3.3 Observations: y and S_ϵ

The vector y is composed of the TBs from the MWR and the mixing ratio and/or temperature from the RL. Its size is variable, since it depends on the number of values the lidar is able to measure in

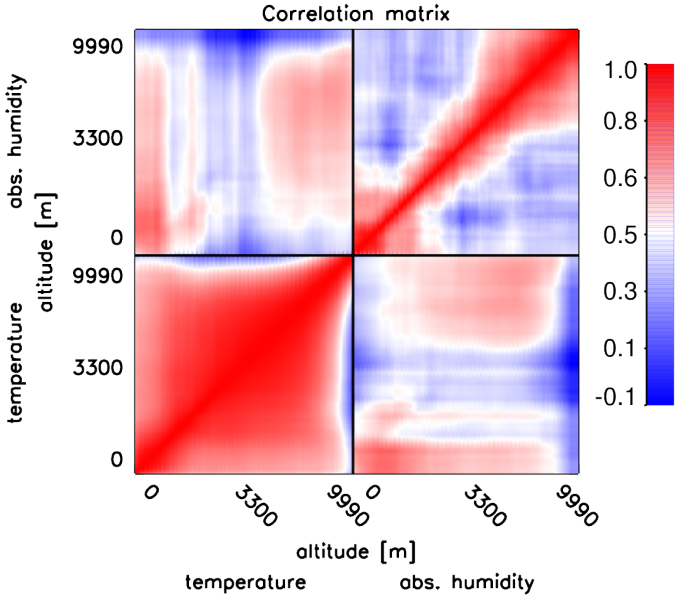

Figure 1. Correlation matrix for the 217 radiosondes in HOPE. Correlation is shown between temperature and absolute humidity as a function of the altitude (from 0 to 10 km). First and fourth quadrants (from up to down and left to right), represent the corr(q,T) and corr(T,q). The second and third, the corr(q,q) and corr(T,T) respectively.

every given profile. A humidity (temperature) lidar profile is provided every 30 meters, from 180 m (from around 1.7 km) to 10 km, with temporal resolution of 5 minutes. The units of these observations are kg/kg (K). For every lidar profile one must determine the range of altitudes where the data can be considered meaningful. This range has been defined via the relative error. The relative error is calculated at each altitude as the ratio between the error and the measurement, as a percentage. When this value is larger than 100%, the data is considered too noisy and is discarded. Care is needed when defining this threshold, because possible random peaks in the error can lead to a missidentification. Therefore, before the analysis, a running average is performed on the data (we choose 300 m window size) previously to the analysis. In general, the 100% error altitude might be reached at different points depending on the weather situation or night/day-times. Typically for water vapor it was found at around 3-4 km during daytime measurements; and around 7-8 km in nighttime measurements.

In the simplest case when one single atmospheric parameter is retrieved, y is composed of $t + m$ elements, where m is the number of altitudes where the lidar measurements have sufficient signal to noise ratio, and t is the number of TBs. Seven brightness temperatures are used for the retrieval of absolute humidity, while 7+20 TBs are used in the case of the temperature retrieval, due to the inclusion of angular information (see Section 2.2). In the case of the simultaneous humidity and

temperature retrieval, the vector y becomes larger and it is formed by $t_q + t_T + m_q + m_T$, that is: the seven TBs of the K-band (t_q), the 27 TBs in the V-band (t_T), the number of valid altitudes for the lidar mixing ratio (m_q) and the number of valid altitudes for the lidar temperature (m_T).

Note that TB can be considered from the MWR directly, while the lidar products (mixing ratio and temperature profiles) are used instead. This is because the lidar raw data requires a complex processing and a clear forward model cannot be defined, see section 2.1 for processing details.

On the one hand, the covariance matrix associated with the MWR measurements was obtained empirically by calculating the correlation between the different channels, while constantly viewing an ambient black-body target with known temperature. It is a 7x7 square matrix for each band. If temperature and humidity are retrieved together, then it becomes a 14x14 matrix. The diagonal elements represent the autocorrelation of each channel, typically with values around the noise level (~ 0.25 K). The off-diagonal elements represent the correlation between the measurements of different channels. Because the channels share some electronics inside the instrument, the off-diagonal correlations cannot be consiered zero, but they typically have values one order of magnitude smaller than the main diagonal.

On the other hand, the part of S_ϵ corresponding to the RL is defined as a diagonal matrix containing the variances of every altitude. This definition implies no correlation between measurements in different heights.

3.4 Forward models (FM)

The forward models for the lidar are trivial, since we are not dealing with raw data, but directly with the products. So the lidar FM for water vapor simply performs the conversion from absolute humidity to mixing ratio or scales the temperature grid. In the case of the temperature, the FM is the unity. The FM for the MWR is more complex since it involves a radiative transfer model (Löhnert et al., 2004). It considers emission and absorption of radiation by gases in the atmosphere but neglects scattering, which can be ignored for all atmospheric particles except for rain droplets. The model divides the atmosphere in layers and calculates the optical thickness and absorption coefficients. From these values, and applying the radiative transfer equation (9) (Janssen, 1993), the TBs are calculated:

$$T_{B,ground} = T_{B,cos} exp(-\tau) + \int_0^\infty T(s)\alpha(s) exp(-\int_0^s \alpha(s')ds')ds \tag{9}$$

Where τ is the optical depth of the whole atmospheric column (opacity), α is the absorption coefficient $[m^{-1}]$ and $T_{B,cos}$ is the cosmic background radiation (approx. 2.7 K) (Janssen, 1993).

The retrieval vertical grid is defined for every profile. It varies, as well as the observation vector, depending on the amount of available lidar information for every given profile. In the atmospheric regions where lidar data is available, the vertical resolution of the retrieval product is 30 meters

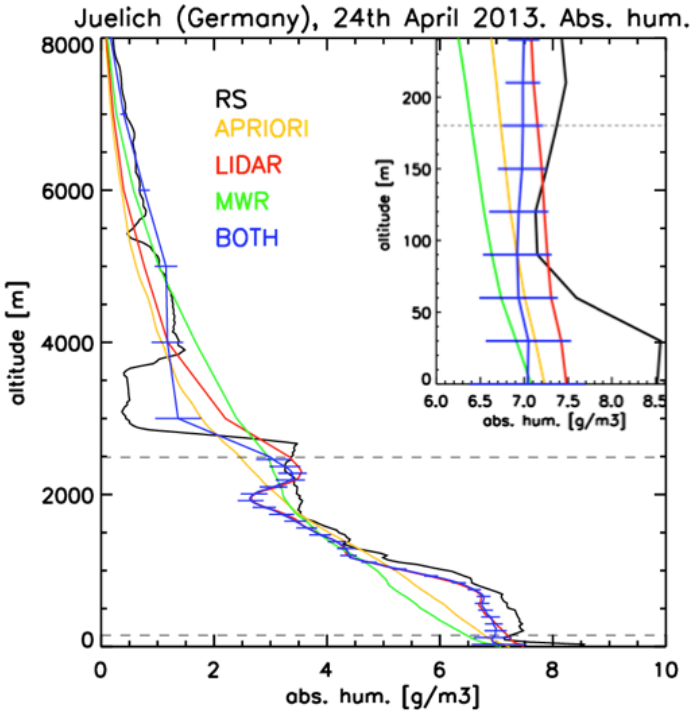

Figure 2. Algorithm performance for a single water vapor profile. Comparison between different instruments: in black, the RS is taken as reference. Yellow is the a priori information. Red is the result of the algorithm with RL only as input. Green is the resulting profile for only MWR. Blue is the combination of both instruments (RL+MWR) with the error bars associated to the retrieved profile. The dashed horizontal lines enclose the region where the lidar data is used. The upper right pannel is a zoom for the region close to the ground, between 0 and 250 m.

(same as the lidar). Above the point where the RL signal is lost, and since the MWR cannot provide more resolution, the algorithm will retrieve one point every 1 km.

4 Absolute humidity retrieval

4.1 Single profile and time series

In a first approach, the OEM has been implemented for the combination of the two instruments to retrieve atmospheric absolute humidity. In addition, it allows to work with a single instrument. This aspect will be interesting to compare the performance of each sensor working alone, with the combination of both.

In the following, the results for one example profile are presented (Fig. 2). The radiosonde launched at 11 UTC on the 24th of April is shown as reference. The a priori profile is the prior atmospheric knowledge, and the starting point for the algorithm.

At first, we introduce in the OEM only the portion of profile where RL data is valid (i.e. from 180 m to 2.5 km, ~ 44layers), not taking into account the MWR. The result of the algorithm is a complete profile from ground to 10 km. In the region with lidar availability, the result will tend to the portion of lidar profile, since the error associated to this measurements is very small (on the order of $0.5g/m^3$). In the regions where no lidar data can be defined, the profile will be completed with the information provided by the a priori profile, which is the only information available.

On the one hand, if only the seven TBs of the MWR are introduced in the OEM, a very smooth profile is obtained. This is because the seven frequencies do not provide enough information to distinguish fine vertical structures: MWR can only provide ~ 2 DOF per profile, as already mentioned in section 2.1. The a priori profile plays a dominant role.

On the other hand, when RL and MWR are combined in the algorithm, the resulting profile is very similar to the part of RL profile in the region from 180 m to 2.5 km. This is again due to the small error associated to the lidar measurements. Outside this region, the profile is completed based on the information provided by the TBs. The theoretical uncertainty of the product is provided by the algorithm as well. The error is small in the region where there is RL data availability ($\sim 0.5g/m^3$), but it increases with altitude, as expected. It is also slightly larger close to the ground ($\sim 1g/m^3$). Similarly, error bars for the only-RL and only-MWR profiles were obtained in the calculations, with larger values than the joint retrieval error, but these are not plotted for the sake of clarity.

The profile obtained with the RL-MWR combination best fits the RS, shown as reference: it is the only case that can detect the drop in humidity at 3 km and the increase at 5 km. It is interesting to pay attention to the lower part of the atmosphere, close to the ground. In figure 2, a zoom from 0 to 250 meters is shown. One can see that the lowest values of the RS are $1-1.5g/m^3$ more humid than the rest of the profiles. This might be explained because the sonde has been launched under different local conditions: while the instruments site is located inside the research center, the RS is launched in an open field area. It could cause slight differencies in the retrieval close to the ground, but should not be a problem in the free troposphere.

At ground level, the two only available sources of information are the MWR and the a priori, which has a much larger uncertainty than the instrument and thus a smaller weight on the result. Thus, the RL-MWR combination tends to the MWR values close to the ground, but quickly approaches to the lidar, as soon as the first RL values are available. The same procedure can be applied, not only to one single profile, but also to a larger measurement period. The result of the combined retrieval is shown in Figure 3, which presents the time series of the absolute humidity on the 17th of April 2013, during HOPE. The figure shows a more humid layer close to the ground with values around $8-9g/m^3$. Fine structures and their temporal evolution are well captured, associated with a cold front.

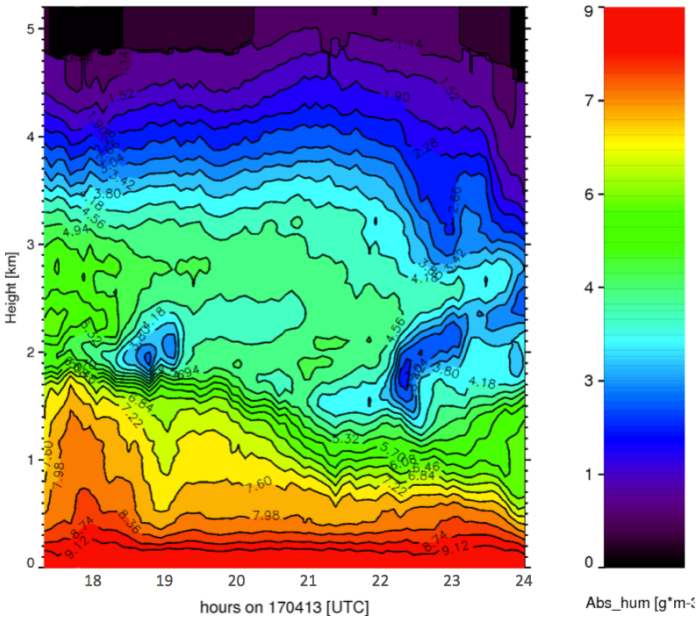

Figure 3. Time series of the joint retrieved absolute humidity, in the afternoon of the 17th April 2013.

4.2 Statistics over HOPE

The absolute humidity algorithm has been applied to all the clear sky periods with simultaneous availability of MWR and RL. The MWR was working continuously, so this selection is restricted to lidar availability. There are in total 4201 lidar profiles (30% of the total campaign). Out of them, 717 profiles have been considered as clear sky (around 17% of the total). Out of all the clear sky profiles, the convergence of the OEM is found in 95.8% of the cases, that is, 687 profiles. In the rest of the cases, the convergence is not found because the algorithm cannot find a profile which is simultaneously consistent with the measurements of the two instruments and the a priori, within their uncertainties.

4.2.1 Integrated Water Vapor

An important parameter to study is the IWV. The measurements of IWV from the Global Position Satellite (GPS) (Bevis et al., 1992) can be used as comparison. In figure 4, the time series of the IWV during HOPE is presented. The continuous IWV signal from the GPS is shown together with the IWV from the joint retrieval, which is only available during clear sky events. Bias and standard deviation are also calculated and shown in Table 1. They are calculated not only for the OEM and the GPS, but also for OEM and a MWR multi-variable regression based IWV retrieval (Steinke et al., 2014). The agreement is very good: in both cases the bias is smaller than $0.6 kg/m^2$ and the standard deviation smaller than $1.2 kg/m^2$. These values lie inside the GPS uncertainty of $1-2 kg/m^2$ (Gendt

Sources	Mean	STD
a) OEM and GPS	-0.288	1.205
b) OEM and MWR	0.599	0.656

Table 1. Mean and standard deviation between the OEM product and a) GPS and b) standard product of the MWR retrieved with multi-variable regression. Units: kg/m^2.

et al., 2004) and the MWR product of $\sim 0.5 - 1 kg/m^2$ (Steinke et al., 2014). This result gives us confidence that the developed OEM method delivers reliable water vapor profiles.

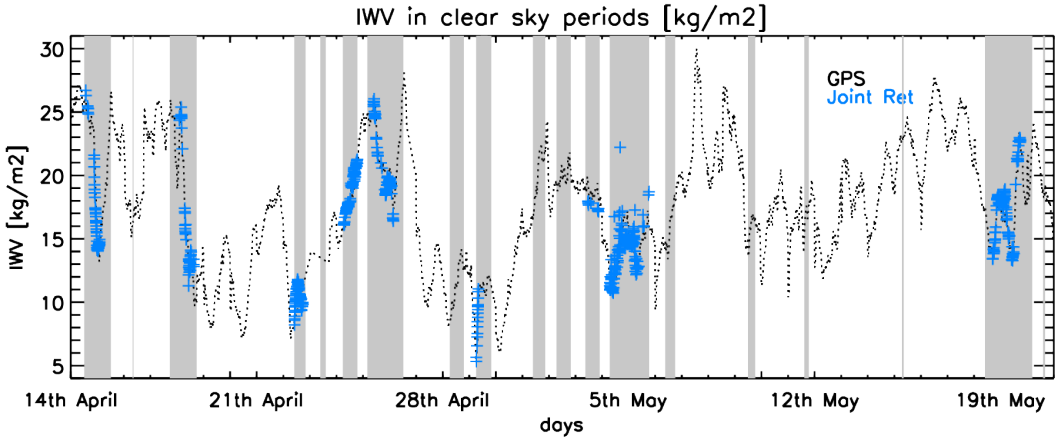

Figure 4. Time series of IWV during the whole HOPE period for clear sky cases. In black: the GPS signal; in blue: the IWV calculated from the joint retrieval (only in clear sky cases). Shaded areas represent the RL availability.

4.2.2 Comparison to RS

As already explained at the beginning of the current section, the retrieval grid of every profile depends on how much data from RL can be taken into account, which will depend on the atmospheric conditions, background noise, etc. Nevertheless, in order to clearly assess the benefits of the sensor synergy, a different retrieval strategy is used for the subsequent tests: the algorithm is run cutting the RL profiles at a fixed altitude to retrieve all the products using the same vertical grid. In this manner, all the RL profiles have been *artificially* cut at an altitude of 2.5 km. In the case that a lidar profile gets too noisy before this altitude, it is discarded and not taken into account for the statistics. This cut-altitude is chosen in order to keep at least 75% of the profiles in the statistics (only 23% of the considered RL profiles reach 100% relative error ar a height lower than 2.5 km). In this way, three regions are defined:

- Region a) from ground to 180 m : no lidar data

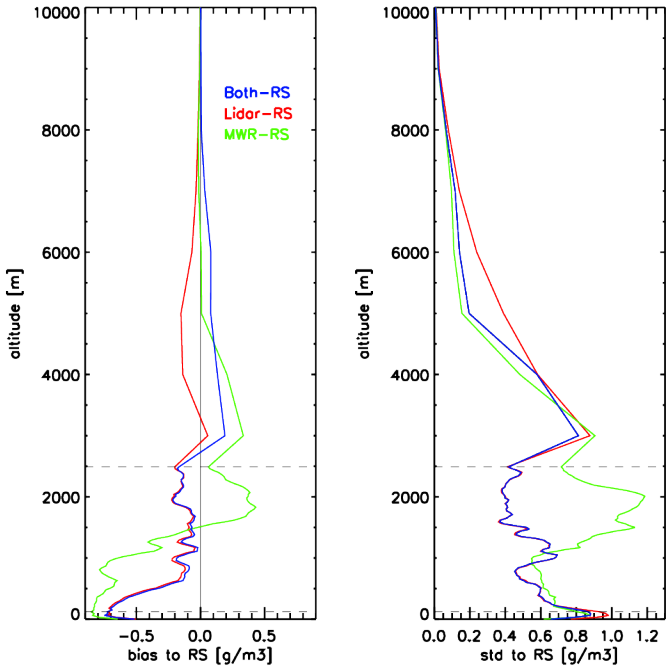

Figure 5. Mean and standard deviation of the difference between the 18 clear sky radiosondes: MWR (in green), RL (in red) and the combination of both (blue). The dashed horizontal lines enclose the region where the lidar data is used.

- Region b) from 180 m to 2.5 km: the only domain where lidar data is available. It is enclosed inside the dashed horizontal lines in figure 5.
- Region c) from 2.5 km to 10 km: no lidar data.

A comparison of the vertical absolute humidity profiles with the radiosonde profiles is performed. In total, 18 valid clear sky radiosondes have been found during the periods where BASIL was measuring. In figure 5, the bias (on the left) and the standard deviation (on the right) to the RS are presented for the three cases: only-MWR, only-RL and the MWR+RL combination.

The region (a), exhibits the largest standard deviations (std) and biases, with similar values for the three cases. That can be explained due to the distance and different environmental conditions where the RS is launched, with respect to the site where the instruments measured.

In region (b), the biases and standard deviation for the only-RL and RL+MWR are very similar, whereby only-MWR reveals the largest values. The similarity between only-RL and the combination is again explained by the small error associated to the lidar measurements: the product of the combination tends to the lidar data when available, as seen in section 4.1. From $\sim 500m$ to 2.5 km, both only-RL and RL+MWR show a small bias on the order of $\sim 0.2g/m^3$, but below this altitude and until the end of region (b), the deviation is increasing up to $\sim 0.75g/m^3$. This fact may suggest

that the lidar data in the lower 500 m could have some problems with the OVF. This feature will be examined in more detail in subsection 4.2.4.

In region (c) all the three values for the different retrievals are similar. The only-MWR seems to perform best when comparing to the RS, because both its bias and stv are the smallest. The only-RL case presents the largest bias and stv because in this region only information from the a priori is provided. The combination of the two sensors presents intermediate values.

Unfortunately, a set of only 18 radiosondes is not enough to asses the benefits of the synergy. In addition, when interpreting the results in figure 5, one must take into account that the RS itself presents some sources of error which are not easy to quantify: launch distance of 4 km to the site, drifting of the balloon, dry bias, etc. Another quantity with the capability to show the improvements of the RL+MWR combination is the theoretical error of the retrieved profiles. This parameter is studied in the following subsections.

4.2.3 Theoretical error comparison

As already mentioned in section 3, the algorithm provides an estimation of the error for the retrievals, see eq. (4). This theoretical error is computed for every profile and for the three different cases: using only-RL, only MWR and the RL+MWR combination.

Figure 6 presents the a priori uncertainty, as well as an average over the 636 theoretical error profiles calculated after running the OEM for all the HOPE clear sky periods. Clearly the uncertainty associated to the a priori is the largest, as it represents the atmospheric variability within the HOPE period. When only the TBs of the MWR are introduced in the algorithm, the average error estimate is reduced at least by half throughout the whole atmosphere with respect to the a priori error. That is possible thanks to the pieces of information introduced by the MWR. When only the lidar information is used to run the algorithm, the error inside region (b) gets much smaller than in the other two previous cases. Compared to the only-MWR error, which has an average of $\sim 0.7 g/m^3$, the only-RL uncertainty is reduced to almost $1/7th$ of this value. In regions (a) and (c) the only-RL error is larger than in region (b) because no lidar data is available and thus the information used to fill the profile is completed with the a priori. The only-RL uncertainty is indeed especially large above 3km, where it tends to the a priori uncertainty, presenting even larger values than the only-MWR error.

However, when the combination of RL and MWR is performed, the achieved error is the smallest for all the altitudes. In region (b), the error is almost the same than for the only-RL case. But outside this region, the MWR contribution plays an important role to reduce the uncertainty. In region (c), from averages uncertainty values of 0.17 and 0.22 g/m^2 for only-MWR and only-RL respectively, the uncertainty of the combination is reduced to an average value of 0.12 g/m^2. Similarly, in the lowest region, the average error for the combination is 0.30, in comparison with 0.71 and 0.33 g/m^2 for the only-RL and only-MWR cases, respectively.

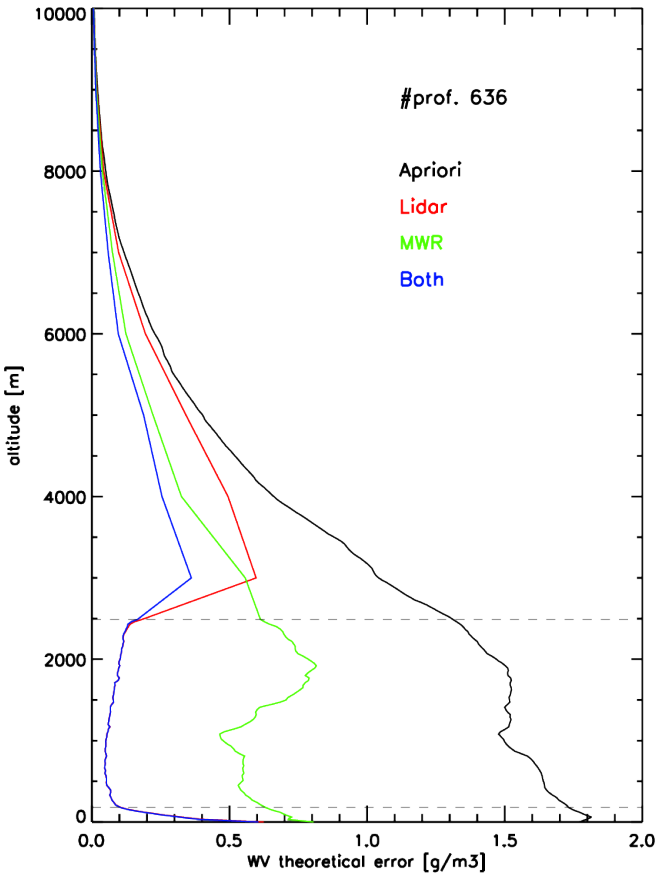

Figure 6. Mean theoretical uncertainty over the 636 clear sky cases during the complete HOPE period. In black: a priori uncertainty. Red: only lidar has been introduced in the algorithm. Green: only MWR. In blue, the combination of the both instruments. The dashed horizontal lines enclose the region where the lidar data is used.

In conclusion, there is an obvious improvement in the theoretical error due to the synergy of the two instruments. One can quantify the relative reduction between the averaged single-instrument and joint theoretical error profiles by dividing the difference among error profiles by the single-instrument one. That way, the absolute humidity error can be reduced in the complete atmospheric profile by 59.8% and 37.9% on average, with respect to the retrieval using only MWR data or only RL, respectively. This improvement is especially clear in region (c), where lidar data are not available. The improvement of the combination in region (a) is better analysed with the experiment in the next subsection.

4.2.4 Sensitivity study in the lower atmosphere

As argumented in section 4.2.2, the high bias values for only-RL and RL+MWR from ground to 500 m (see figure 5), might reveal a problem with the lidar OVF in this region. In order to exclude this possible problem and to show the good performance of the MWR in the lowest layers of the atmosphere, another experiment is performed. Here, we will assume that the OVF of the RL does not allow us to receive valid measurements from the lowest 500 m. Under this condition, the lidar data from 180 to 500 meters is discarded for all the profiles. The algorithm is run again for the complete HOPE period taking this condition into account.

The results are shown in figure 7, together with the initial OVF starting at 180 meters. In both cases (regular OVF and increased OVF), the results are very similar when the RL data is available (from 500 m to 2.5 km). But in the lower region for the case of the increased OVF, the combination of the two instruments is clearly better: there is an uncertainty reduction at the ground level of about $0.1 g/m^3$ from the combination with respect to the only-RL, which is gradually reduced towards the total lidar overlap until $\sim 400m$. This confirms that the MWR contributes with higher information content in the lower atmosphere. Above this point and up to 2.5 km, the error is almost equal for the cases of regular OVF and increased OVF. From 2.5 km to 10 km, the increase of the OVF shows a slight increase in the theoretical error of $\sim 0.05 g/m^3$ and $\sim 0.02 g/m^3$for the RL+MWR and only-RL cases, with respect to the regular OVF.

4.2.5 Increase of the RL error

In section 3.3 the components of the covariance matrix S_e were determined to our best knowledge. However, it might be possible that additional uncertainty sources exist. When we compared the theoretical uncertainty for the different instument configurations (Fig. 6), the only-RL error at 2 km was $0.1 g/m^3$ which is significantly lower than the deviation with respect to the RS at the same altitude ($0.4 g/m^3$; Fig.5). That means that the theoretical error is about four times smaller than the standard deviation to the RS. This fact could suggest that the error associated to the lidar is very small, or in other words: that the initial lidar uncertainty was not properly defined. As explained in section 3.3, only Poisson noise was taken into account but there can be other possible sources of uncertainty. For this reason, we artificially incremented the RL error by a factor of 4 to study the sensitivity of the retrieved profile with respect to the RL measurement uncertainty.

The results of this test are plotted in figure 7, together with the initial values (without increment), for the only-lidar and combination cases. The new averaged errors have a very similar starting point at the ground, but they have increased by a factor of 2 to 3 in region (b). The difference between the increased errors in the only-RL and RL+MWR cases is more noticeable than the original cases (with no increment), especially from 2 km upwards.

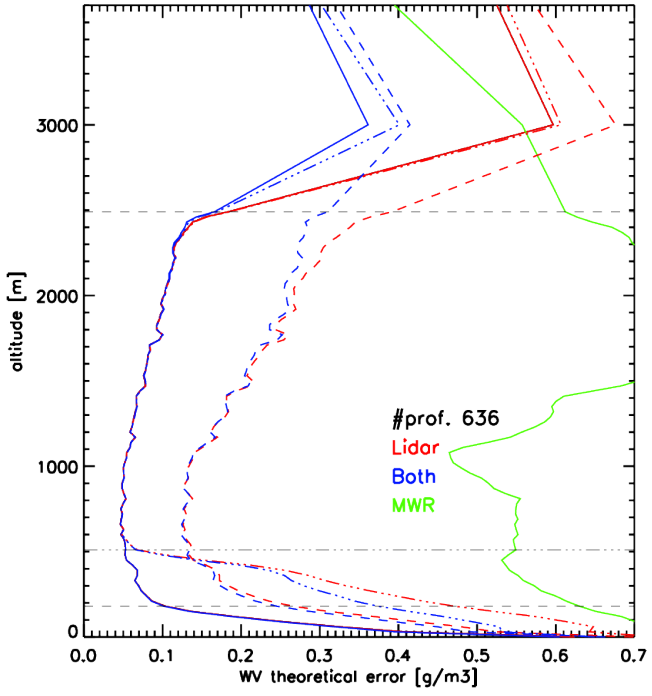

Figure 7. Mean theoretical error over the 636 clear sky cases during the complete HOPE period. Red: only RL has been introduced in the algorithm. Green: only-MWR. In blue, the combination of RL and MWR. The dashed horizontal black lines define the region where lidar data has been considered available. The dashed red and blue lines represent the result when the lidar error has been incremented by a factor of four. The dotted-dashed red and blue lines correspond to the case where lidar data has been suppresed from ground until 500 meters. Solid lines show the errors without increments, as shown in figure 6.

A change is also observed on the averaged DOF of the profiles (table 2), which allows to study the amount of information provided by the different instruments in the three different atmospheric regions. Table 2 summarizes these mean values in the different regions.

For the only-RL case: in the regions where no lidar data is available ((a) and (c)), the DOF are as expected, zero. But in region (b), the DOF are very high, meaning that the instrument provides a large information content, indirectly explained because the error of the lidar is small. The MWR-only presents a much smaller number of DOF but distributed in the whole profile. Proportionally, the DOF are higher for lower altitudes, which confirms the better performance of the MWR close to the ground. The numbers for the MWR+RL combination show that, thanks to the inclusion of MWR, the DOF in regions (a) and (c) are not zero any more and still in region (b) the DOF remain almost the same. In any case, the total average number of DOF in the column is largest for the combination of the two instruments, increasing in almost 2 DOF with respect to the only-RL case, and in almost 25 DOF with respect to the only-MWR profile. That is another benefit of the synergy.

Region	RL	MWR	Combination
a) Ground to 180 m	0.00	0.07	0.03
b) 180 m to 2.5 km	25.90	1.01	25.75
c) 2.5 km to 10 km	0.00	1.18	1.69
Total	25.90	2.26	27.47
Region	RL	MWR	Combination
a) Ground to 180 m	0.00	0.07	0.06
b) 180 m to 2.5 km	12.19	1.01	12.11
c) 2.5 km to 10 km	0.00	1.18	1.57
Total	12.19	2.26	13.74

Table 2. Degrees of freedom for signal comparison for absolute humidity. Average over 636 profiles. The atmosphere is separated in three regions according to lidar availability. The DOF are presented for three cases: only RL, only MWR and the combination of both instruments. In the upper part, no increment on the RL error has been considered. In the bottom part, the RL uncertainty has been multiplied by a factor of four.

When an increment in the RL uncertainty is considered, the amount of useful information provided by this instrument is smaller, and thus the DOF are reduced. This reduction is experimented in all the regions where the RL is involved. The numbers for the MWR only retrieval, remain the same, because no change on this instrument is done.

To help in the interpretation of the numbers in table 2, figure 8 has been included. This figure represents the cumulative degrees of freedom per profile for the different instrument combinations. In the case of only-MWR, the cumulative DOF are smaller than for the other cases, reaching a maximum of about 2. Whenever lidar is available, the DOF increase linearly, thanks to the strong lidar information content. In the case of only-RL, above 2.5 km, the cumulative DOF remain constant because no additional information is introduced. Nevertheless, for the RL+MWR, the cumulative DOF is still increasing above 2.5 km thanks to the inclusion of the MWR information.

The results presented so far confirm that the RL+MWR water vapor synergy is meaningful and successful. In addition, they suggest that a careful specification of the instrument errors, specially for the RL, is required.

5 Temperature retrieval

The OEM has been used so far to retrieve atmospheric absolute humidity profiles by combining RL and MWR. In addition, this method can be applied to the retrieval of temperature profiles. Nevertheless, due to the restricted temperature data availability for the RL, no long term statistics are analyzed. Therefore, a single example profile is presented to illustrate the capabilities of the OEM applied to temperature. Similarly to the water vapor retrieval, the OEM allows to work with one single instrument or with the combination of both. The angular information of the brightness tem-

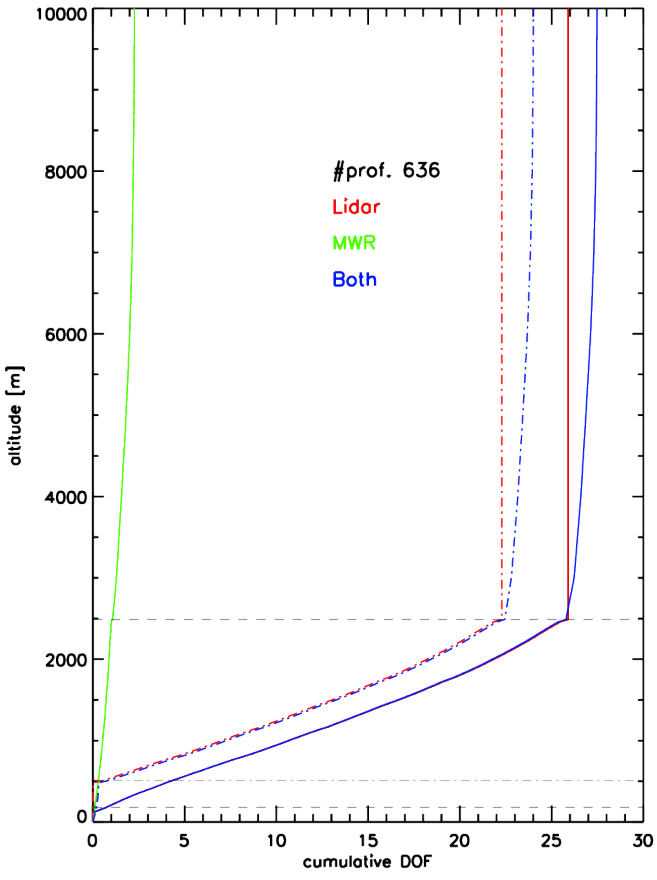

Figure 8. Cumulative degrees of freedom per profile for the different instrument combinations: in red, only-RL; in green, only-MWR and in blue, the combination of the two sensors. The dotted-dashed lines represent the degrees of freedom for the case where the overlapping function has been extended up to 500 m. The average number of DOF in every region are summarized on Table 2. The dashed horizontal grey lines enclose the part of the atmosphere where lidar data has been considered.

peratures along the 60 GHz oxygen absorption complex is included, which can improve the MWR temperature retrievals (Crewell and Lohnert, 2007). A scanning strategy over six angles is defined for the instrument: 90, 42, 30, 19.2, 10.2 and 5.4 degrees. As already shown in Crewell and Lohnert (2007), the best results in the retrievals are achieved when considering the four most opaque frequencies with their angular information and the three more transparent channels with only their zenith measurements. That implies that the observation vector is enlarged introducing 20 more values for angular information. Figure 9 shows the retrieved profiles and their deviations to the radiosonde on the 17th of April 2014, at 23 UTC. Again, three atmospheric regions are differentiated according to the lidar availability in this particular profile:

- Region a) from ground to 2.5 km, where the RL error is large due to OVF problems.

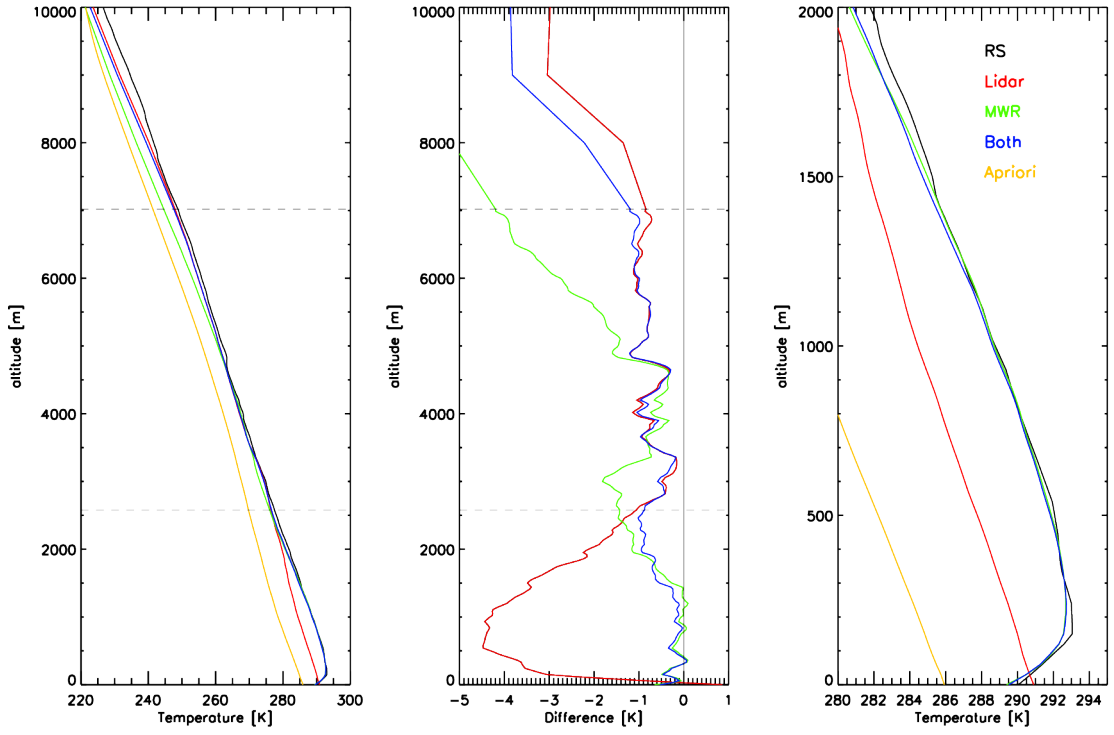

Figure 9. Example profile for temperature retrieval, the 17th of April 2014, at 23 UTC. (a) Complete profiles of temperature for (black) the radiosonde, (red) only-RL information, (green) only MWR, (blue) the combination of MWR and RL. The horizontal dashed grey lines enclose the area where RL data was available. (b) Difference with respect to the radiosonde and (c) zoom to the lower 2km of the atmosphere.

- Region b) from 2.5 km to 7 km, the only domain where lidar data can be considered valid.
- Region c) from 7 km to 10 km, where the RL signal gets too noisy.

In the case of the temperature, the lidar profile for this specific case study is much more affected by the OVF than in the water vapor profile, and so there are no valid temperature lidar measurements under 2.5 km. The resulting profiles are compared to the RS. In a first approach, the algorithm is run with RL-only data. The resulting profile has a large error in region (a) where the difference to the RS reaches values larger than 4 K. This is because the result tends to the a priori information. In region (b) the difference is reduced to values smaller than 1 K.

In a second step, the OEM is also run introducing only the TBs of the MWR, taking into account the angular information. In this case, an inversion of the temperature close to the ground is detected, which cannot be resolved by the lidar, see right panel on figure 9. The only-MWR performs better in region (a), reducing the difference to more than one fourth of the only-RL value in the lowest 1.5 km. The deviation with respect to the radiosonde grows with the altitude, taking on larger values in region (b) from 5 km above.

Region	RL	MWR	Joint
a) 0 m to 2.5 km	0.00	2.64	2.57
b) 2.5 km to 7 km	8.15	0.47	6.60
c) 7 km to 10 km	0.00	0.07	0.05
Total	8.15	3.19	9.23

Table 3. Degrees of freedom for temperature retrieval, separated in three regions in the atmosphere. Lidar data is only present in region (b). The DOF are presented for the cases were only-RL is used, only-MWR and for the combination of the both instruments.

When the combination RL+MWR is performed, the result is strongly improved when it is compared to the reference. It presents the smallest deviation to the RS in regions (a) and (b), presenting deviations of around $\sim 1K$) up to 7 km. In the lower 2 km of the atmosphere there is a strong improvement of the joint RL+MWR retrieval, because the MWR performs better in this region and the angular scanning is able to enhance the information content. To sum up, it is shown that the total profile error is reduced by 47.1% and 24.6% with respect to the only-MWR and only-RL profiles, respectively.

The degrees of freedom for the temperature profiles are also presented in table 3. The independent pieces of information are improved in the lower part of the atmosphere when introducing MWR information. The combination RL-MWR presents the highest information content, increasing the number of DOF in more than one, with respect to the only-RL case, and in ~ 6, with respect to the only-MWR profile.

6 Simultaneous absoulte humidity and temperature retrieval: relative humidity

Including joint information on water vapor and temperature should lead to improvements on the RH estimates which are of particular interest to study cloud formation. In section 3.2, the correlation information among T and q as a function of the altitude was presented (Fig. 1). In this section, the OEM has been also implemented to retrieve temperature, absolute humidity and relative humidity simultaneously, taking into account that all these three parameters are not independent. The results of running the simultaneous T-q algorithm for RL+MWR, are shown in figure 10, and are also compared to the individual profiles obtained separately as described in Sections 4 and 5. In these two cases, the resulting T and AH profiles are very similar (see figure 10) and no remarkable changes are evidenced. But the RH profiles present some differences. Even if in the lower 5 km the two profiles are alike, above this altitude the resulting RH profile which is calculated introducing T-q correlation, presents a $\sim 20\%$ smaller deviation to the RS than in the case where T-q are retrieved independently. This is the main advantage of using the T-q correlation.

Unfortunately, the RL temperature product is not always available. Because of this reason, we wanted to investigate wheater the simultaneous retrieval of RH is still reasonably good when only using the RL mixing ratio profiles in the cases where there are no RL temperature data. MWR information is kept the same.

The simultaneous T-q algorithm is run again without taking into account the RL temperature profile. Results are shown in figure 10, being very similar to the case when RL temperature was used. Indeed, the no-temperature RL profile presents the smallest average deviation to the RS ($\sim 2.5\%$) in the complete profile, compared to the case where RL temperature was included ($\sim 4.3\%$) and the case when T and q where retrieved independently ($\sim 7.8\%$).

On the one hand, we have shown that, for a particular case study, the introduction of correlation information T-q is benefitial because it reduces the deviation to the RS, specially in the upper part of the atmosphere, where there is no RL water vapor signal. On the other hand, we demonstrate that the RL temperature information is not essential and that the RH retrieval is still good when this information is omitted. Unfortunately, since the RL temperature data availability is reduced during the campaign, a further investigation with more case studies cannot be carried out at present for HOPE data.

7 Conclusions

Humidity and temperature are essential variables for the description of any meteorological process. Highly resolved, accurate and continuous measurement of these parameters are required for a deeper understanding of many atmospheric phenomena. Unlikely, single instruments available nowadays are not able to provide vertical coverage, vertical and temporal resolution of the humidity and temperature atmospheric profiles. This is the motivation why the synergy of different sensors has become a trend in the last years.

In this paper, a new method to combine Raman lidar and microwave radiometer measurements has been presented. The joint algorithm that combines the two sensors is based on an Optimal Estimation Method. Results for 53 hours of clear sky measurements during the HOPE period are presented for water vapor retrievals, together with one temperature and relative humidity case study.

The improvements of the synergy have been analysed in terms of several parameters, like the reduction of the theoretical error or the increase of DOF, showing strong advantages with respect to the two instruments working separately. For example, when applying the combined retrieval to the complete HOPE period, the absolute humidity error can be reduced by 59.8% and 37.9% on average, with respect to the retrieval using only MWR data or only RL, respectively. Results for a case study temperature profile show that the error is improved in a 47.1% and 24.6% with respect to the only-MWR and only-RL profiles, respectively. The synergy present its strongest advantages

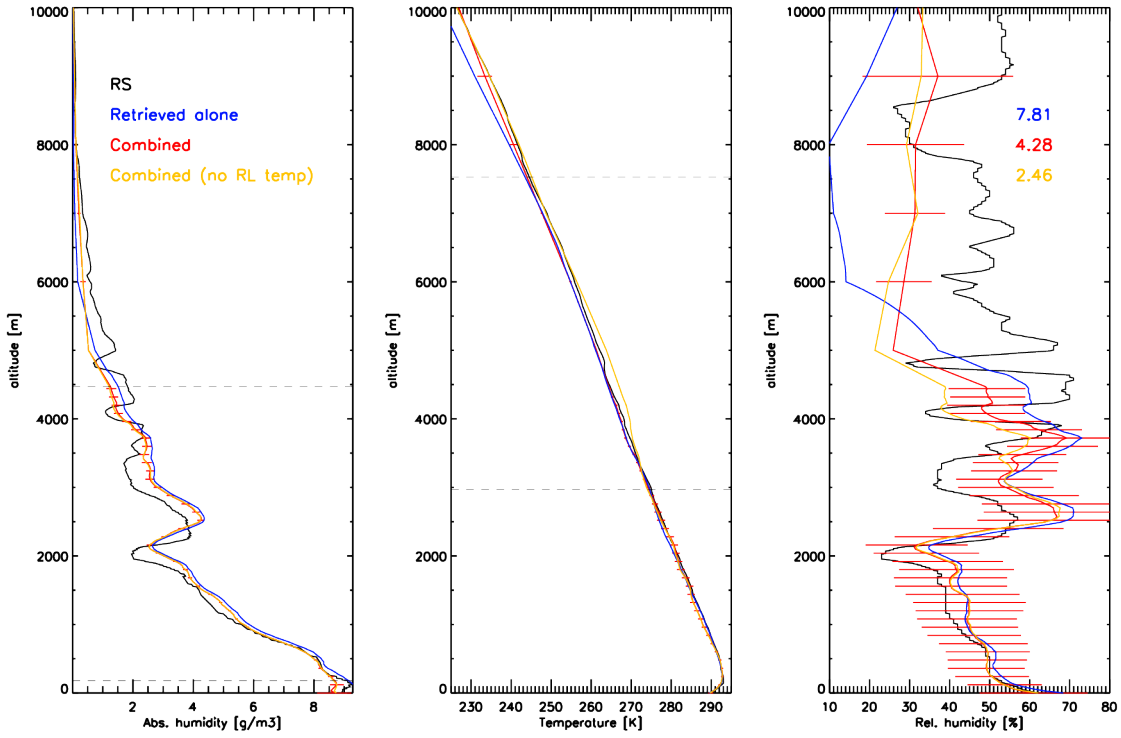

Figure 10. Absolute humidity, temperature and relative humidity from RS (black), profiles retrieved separately using MWR+RL (blue), the simultaneous T-q retrieval using MWR+RL(red) and the simultaneous T-q retrieval without RL temperature (yellow). Horizontal bars represent the error associated to the resulting profiles. The horizontal grey dashed lines enclose the areea where lidar data was available. Numbers represent the averaged difference to the RH of the RS for each case in percentage [%].

in the regions where RL data is not available, whereas in the regions where both instruments are available, RL dominates the retrieval.

One relative humidity profile has been retrieved by assuming temperature and humidity correlations in the atmosphere, calculaded from RS data. The joint information on T-q leads to improve the RH estimates, which will be of particular interest to study cloud formation. In addition, it has been shown than the RH profiles can be successfully retrieved without using RL temperature information. A larger data set to study the temperature and relative humidity retrievals could be desiderable.

With the expansion of the ground based network of atmospheric profiling stations the application of the OEM at several sites under different climate conditions will become possible. In this respect, the definition of an appropiate background error covariance needs to be carefully addressed. Further studies will extend the algorithm to cloudy cases. In addition, the method will be applied, not only to ground based measurements, but also to airbone data (Mech et al., 2014), which will allow to complete the study of meteorological phenomena from the airborne point of view.

Acknowledgements. Acknowledgements: This research has been financed by ITARS (www.itars.net), European Union Seventh Framework Programme FP7: People, ITN Marie Sklodowska Curie Actions Programme under grant agreement no 289923. The authors would like to acknowledge the Federal Ministry of Education and Research 25 in Germany (BMBF), who, through the research programme “High Definition Clouds and Precipitation for Climate Prediction" – $HD(CP)^2$, financed HOPE. Special thanks to Kerstin Ebell (for her important contribution to the early stages of the project), Dave Turner (for his always fruitful ideas) and Bjorn Stevens (for his useful advice).

References

Bevis, M., Businger, S., Herring, T. A., Rocken, C., Anthes, R. A., and Ware, R. H.: GPS meteorology: Remote sensing of atmospheric water vapor using the global positioning system, Journal of Geophysical Research: Atmospheres, 97, 15 787–15 801, doi:10.1029/92JD01517, 1992.

Crewell, S. and Lohnert, U.: Accuracy of Boundary Layer Temperature Profiles Retrieved With Multifrequency Multiangle Microwave Radiometry, Geoscience and Remote Sensing, IEEE Transactions on, 45, 2195–2201, doi:10.1109/TGRS.2006.888434, 2007.

Di Girolamo, P., Marchese, R., Whiteman, D. N., and Demoz, B. B.: Rotational Raman Lidar measurements of atmospheric temperature in the UV, Geophysical Research Letters, 31, n/a–n/a, doi:10.1029/2003GL018342, 2004.

Di Girolamo, P., Summa, D., Bhawar, R., Di Iorio, T., Cacciani, M., Veselovskii, I., Dubovik, O., and Kolgotin, A.: Raman lidar observations of a Saharan dust outbreak event: Characterization of the dust optical properties and determination of particle size and microphysical parameters, Atmospheric Environment, 50, 66–78, doi:10.1016/j.atmosenv.2011.12.061, 2012.

Ebell, K., Orlandi, E., Hünerbein, A., Löhnert, U., and Crewell, S.: Combining ground-based with satellite-based measurements in the atmospheric state retrieval: Assessment of the information content, Journal of Geophysical Research: Atmospheres, 118, 6940–6956, doi:10.1002/jgrd.50548, 2013.

Gendt, G., Dick, G., Reigber, C., Tomassini, M., Liu, Y., and Ramatschi, M.: Near Real Time GPS Water Vapor Monitoring for Numerical Weather Prediction in Germany, Journal of the Meteorological Society of Japan. Ser. II, 82, 361–370, 2004.

Girolamo, P. D., Summa, D., and Ferretti, R.: Multiparameter Raman Lidar Measurements for the Characterization of a Dry Stratospheric Intrusion Event, 26, 2009.

Han, Y., Westwater, E. R., and Ferrare, R. A.: Applications of Kalman Filtering to Derive Water Vapor Profiles from Raman Lidar and Microwave Radiometers, Journal of Atmospheric and Oceanic Technology, 14, 480–487, doi:10.1175/1520-0426(1997)014<0480:AOKFTD>2.0.CO;2, 1997.

Janssen, M.: Atmospheric remote sensing by microwave radiometry, Wiley-interscience publication, Wiley, 1993.

Löhnert, U. and Maier, O.: Operational profiling of temperature using ground-based microwave radiometry at Payerne: prospects and challenges, Atmospheric Measurement Techniques Meas. Tech., 2012.

Löhnert, U., Crewell, S., Simmer, C., and Macke, A.: Profiling cloud liquid water by combining active and passive microwave measurements with cloud model statistics, Journal of Atmospheric and Oceanic Technology, 18, 1354–1366, 2001.

Löhnert, U., Crewell, S., and Simmer, C.: An Integrated Approach toward Retrieving Physically Consistent Profiles of Temperature, Humidity, and Cloud Liquid Water, Journal of Applied Meteorology, 43, 1295–1307, doi:10.1175/1520-0450(2004)043<1295:AIATRP>2.0.CO;2, 2004.

Löhnert, U., van Meijgaard, E., Baltink, H. K., Groß, S., and Boers, R.: Accuracy assessment of an integrated profiling technique for operationally deriving profiles of temperature, humidity, and cloud liquid water, Journal of Geophysical Research: Atmospheres, 112, n/a–n/a, doi:10.1029/2006JD007379, 2007.

Löhnert, U., Turner, D. D., and Crewell, S.: Ground-Based Temperature and Humidity Profiling Using Spectral Infrared and Microwave Observations. Part I: Simulated Retrieval Performance in Clear-Sky Conditions, Journal of Applied Meteorology and Climatology, 48, 1017–1032, doi:10.1175/2008JAMC2060.1, 2009.

Löhnert, U., Schween, J. H., Acquistapace, C., Ebell, K., Maahn, M., Barrera-Verdejo, M., Hirsikko, A., Bohn, B., Knaps, A., O'Connor, E., Simmer, C., Wahner, A., and Crewell, S.: JOYCE: Jülich observatory for cloud evolution, B. Am. Meteorol. Soc., 2014.

Maschwitz, G., Löhnert, U., Crewell, S., Rose, T., and Turner, D. D.: Investigation of ground-based microwave radiometer calibration techniques at 530 hPa, Atmospheric Measurement Techniques, 6, 2641–2658, doi:10.5194/amt-6-2641-2013, 2013.

Mech, M., Orlandi, E., Crewell, S., Ament, F., Hirsch, L., Hagen, M., Peters, G., and Stevens, B.: HAMP - the microwave package on the High Altitude and LOng range research aircraft (HALO), Atmospheric Measurement Techniques, 7, 4539–4553, doi:10.5194/amt-7-4539-2014, 2014.

Rodgers, C.: Inverse Methods for Atmospheric Sounding: Theory and Practice, World Scientific Publishing Company, Incorporated, 2000.

Rose, T., Crewell, S., Löhnert, U., and Simmer, C.: A network suitable microwave radiometer for operational monitoring of the cloudy atmosphere, Atmospheric Research, 75, 183 – 200, doi:http://dx.doi.org/10.1016/j.atmosres.2004.12.005, cLIWA-NET: Observation and Modelling of Liquid Water Clouds, 2005.

Schneebeli, M.: Advancements in Ground-Based Microwave Remote Sensing of the Troposphere - Calibration, Data Retrieval and Applications, Ph.D. thesis, 2009.

Stankov, B. B.: Multisensor Retrieval of Atmospheric Properties, Bulletin of the American Meteorological Society, 79, 1835–1854, doi:10.1175/1520-0477(1998)079<1835:MROAP>2.0.CO;2, 1998.

Steinke, S., Eikenberg, S., Löhnert, U., Dick, G., Klocke, D., Di Girolamo, P., and Crewell, S.: Assessment of small-scale integrated water vapour variability during HOPE, Atmospheric Chemistry and Physics Discussions, 14, 22 837–22 879, doi:10.5194/acpd-14-22837-2014, 2014.

Stevens, B. and Bony, S.: Water in the atmosphere, Physics Today, 66, 2013.

Turner, D. D., Clough, S. A., Liljegren, J. C., Clothiaux, E. E., Cady-Pereira, K. E., and Gaustad, K. L.: Retrieving Liquid Water Path and Precipitable Water Vapor From the Atmospheric Radiation Measurement (ARM) Microwave Radiometers, IEEE Transactions on Geoscience and Remote Sensing, 45, 2007.

Whiteman, D. N.: Examination of the traditional Raman lidar technique. I. Evaluating the temperature-dependent lidar equations, Appl. Opt., 42, 2571–2592, doi:10.1364/AO.42.002571, 2003.

Wilks, D. S.: Statistical Method in the Atmospheric Sciences, Interational Geophysics Series, Elsevier, 2006.